5.10 Targeting Pre-mRNA Processing in Cancer

TR Webb, SRI International, Menlo Park, CA, United States
S Stamm, University of Kentucky, Lexington, KY, United States

© 2017 Elsevier Ltd. All rights reserved.

5.10.1 Introduction

The processing of pre-mRNA to mature mRNA in metazoans is one of the numerous critical processes for development and normal functioning of cells that is mediated by ribonucleoproteins (RNPs) (see Fig. 1). The importance of the pre-mRNA splicing process as a source of evolutionary adaptability of metazoans appears to justify the significant burden that is born by each cell that must orchestrate the synthesis and regulation of some 200 proteins and numerous functional RNPs, which are critical for the pre-mRNA splicing process. This process involves the highly regulated combinatorial removal of intervening sequences from pre-mRNA followed by the ligation of exons to form mature mRNA, which is then exported into the cytosol from the nucleus. This splicing process is catalyzed and regulated by a highly complex macromolecular complex called the spliceosome (see Fig. 2). The spliceosome is a unique macromolecular machine in its complexity, given that it has thousands of pre-mRNA substrates and potentially millions of...
spliced, and alternately spliced, mature mRNA products that are produced in the transcriptome of each cell; additionally each of these splicing reactions requires regulation that is controlled by the specifics of the environment of the particular cellular phenotype that is undergoing the process of transcription and splicing.1

The spliceosome is a complex of five small nuclear RNPs (snRNPs) (U1, U2, U4, U5, and U6) and approximately 200 associated proteins.2,3 This pre-mRNA maturation process also includes alternative splicing (AS), which is the mechanism that allows for different forms of mature mRNAs to be generated from the same pre-mRNA. Commonly, AS patterns determine the inclusion or exclusion of portions of the coding sequence in the mRNA, giving rise to protein isoforms that differ in their peptide sequence. Alternate splicing is regulated by numerous spliceosomal trans-acting proteins, which are in turn regulated by cis-acting regulatory sites on the pre-mRNA that are the substrates for the spliceosome.2 Since pre-mRNAs for a given gene may contain many different exon and intron combinations, there are often a very large number of possible mRNAs that can result in a correspondingly large set of resulting proteins that may have very different, even opposing, biological functions within the cell. Numerous genes are subject to splicing events that can be either oncogenic or serve to limit potential tumorigenesis, examples include BCL-2-like protein 1, VEGF-A, Fas Cell Surface Death Receptor (FAS), pyruvate kinase isozymes M1/M2 (PKM), or MDM2 (as described in a recent review).4

Recently, there has been a strong acceleration in spliceosome-related research and in the rate of appearance of high-profile publications on this topic, many of which have a significant impact on oncology drug discovery. A search of Pubmed with the query “spliceosome” shows 92 publications in 2000 but 229 publications for 2015. Some of the major recent accomplishments in this
field include the very recent reports of the cryo-EM structures of spliceosome components that have clarified numerous aspects of the splicing process at near (but not at) atomic resolution. The complexity of the spliceosome and the current lack of high-resolution structures complicate progress in understanding many of the critically important molecular and functional mechanisms associated with pre-mRNA splicing. Since the splicing process is critical to the normal functioning of the cells of higher organisms,

**Fig. 2** Pre-mRNA splicing by the major spliceosome: (A) conserved sequence elements of metazoan and yeast pre-mRNAs. Here, two exons (blue) are separated by an intron (gray). The consensus sequences in metazoans and yeast at the 5' splice site (SS), branch point sequence (BPS), and 3' splice site (SS) are as indicated, where N is any nucleotide, R is a purine, and Y is a pyrimidine. The polypyrimidine tract is a pyrimidine-rich stretch located between the BPS and 3' SS. (B) Cross-intron assembly and disassembly cycle of the major spliceosome. The stepwise interaction of the spliceosomal snRNPs (colored circles), but not non-snRNP proteins, in the removal of an intron from a pre-mRNA containing two exons (blue) is depicted. Only the spliceosomal complexes that can be resolved biochemically in mammalian splicing extracts are shown. Eight evolutionarily conserved DExD/H-type RNA-dependent ATPases/helicases act at specific steps of the splicing cycle to catalyze RNA–RNA rearrangements and RNP remodeling events. These enzymes include Sub2 (UAP56 in humans), Prp5, Prp28, Br2r, Prp2, Prp16, Prp22, and Prp43 (with Br2r and Prp22 acting at more than one step in the cycle). The GTPase Snu114 also functions at several steps during the cycle. In yeast, Prp28 acts at a later stage during spliceosome activation (the B complex to B* complex transition). Several of these proteins, such as Prp5, Prp16, and Prp22, also carry out proofreading functions at the stages where they are shown. (C) Cross-exon splicing complexes form on long introns during the earliest stage of spliceosome assembly. An SR protein containing a serine-arginine-rich (SR) domain and RRM (RNA recognition motif) is depicted as interacting with an exonic splicing enhancer (ESE). The U1 (blue) and U2 (green) spliceosomal snRNPs and the two subunits of the U2 auxiliary factor (U2AF), U2AF65 and U2AF35, are also shown interacting with the splice sites flanking the exon. Reprinted from Wahl, M. C.; Will, C. L.; Luhrmann, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* 2009, 136, 701–718, with permission from Elsevier.
this area will remain an active area of research for the foreseeable future. However the remaining mysteries surrounding the splicing process do not diminish the importance of this fascinating molecular machine, which is now recognized as a major frontier for molecular biology and, more recently, for oncology target discovery. The discovery of new small molecules that modulate splicing will likely be key in the facilitation of a more detailed understanding of this emerging field. The prominence of pre-mRNA splicing in the regulation of gene expression also makes it an attractive and innovative potential target for therapeutic intervention in numerous diseases, including cancer.3,2,6,7 This article will be focus specifically on the numerous developments that have led to the emergence of a new field: spliceosome medicinal chemistry.

5.10.1.1 Pre-mRNA Splicing and the Spliceosome

Nearly all human RNA polymerase II (RNA Pol II) transcripts undergo pre-mRNA splicing, a nuclear process where sequences called exons (exported sequences) are joined together and the intervening sequences (introns) are removed.10 Exons are surrounded by three characteristic sequence elements: the 5′ and 3′ splice sites located at the exon-intron border and the branch point located 18–50 nucleotides (nt) upstream of the 3′ splice site. Two splicing systems operate in humans, a major one where the introns are marked by GU-AG bases and a minor one (accounting for less than 0.5% of introns) marked by AU-AC dinucleotides.11 The splice sites are recognized through interaction with five small nuclear RNAs (snRNAs), the U1, U2, U4, U5, and U6 snRNAs. Through base complementarity, the U1 snRNA recognizes the 5′ splice site and the U2 snRNA binds to the branch point sequence. The U4/U5 and U6 snRNAs form a complex (the U4/U5/U6 tri-snRNP) involved in splicing catalysis. Human splice sites characteristically adhere weakly to consensus sequences, and only the GU-AG dinucleotides are consistently conserved in the major spliceosome. The removal of introns occurs in two steps: in the first step, the first nucleotide of the intron is ligated to an adenosine branch point by a nucleophilic attack of the branchpoint 2′-OH on phosphorus to form a 5′-2′ phosphodiester bond, and in the second step the two exons are ligated through another nucleophilic attack on phosphorus by the exon 3′-OH to form the phosphodiester bond between the 5′ and 3′ exons.

The vast majority (>95%) and possibly all multiexon genes undergo AS, where an exon is either included or skipped in the final pre-mRNA.12,13 On average, a human gene contains a median of 26 exons per gene and generates 3.4 isoforms, indicating that AS increases the coding capacity of the human genome more than threefold. AS is more abundant in humans than in other species, for example, mice generate only 2.4 isoforms per gene and C. elegans 1.2 isoforms.10 The overall purpose of AS is to increase the diversity of the mRNA expressed from the genome.14 A genome-wide analysis showed a large influence of alternative splice variants on protein interaction networks. Within an interaction network, most AS variants behave more like distinct proteins than like minor variants,15 suggesting that AS generates functionally distinct proteins from a single gene. The deregulation of alternative pre-mRNA splicing is a hallmark of cancer16 (Fig. 3).

5.10.1.2 Recognition of Exons

Exons on average constitute less than 10% of a human pre-mRNA. They have to be recognized with high fidelity, as omission of a single nucleotide will generate a frameshift that usually generates aberrant proteins. Because of the poor conservation of the splice sites, additional RNA sequences (cis elements) are needed for their recognition. These cis-elements bind to proteins that can be subdivided into two classes: serine-arginine-rich (SR) proteins and heterogeneous nuclear RNPs (hnRNPs). SR-proteins are RNA-binding proteins characterized by an RNA-binding domain and an RS domain that interacts with other proteins.17 hnRNPs are a diverse group of nuclear proteins that bind to pre-mRNA.18 As a general rule, SR-proteins promote exon inclusion as they bind to components of the spliceosome, whereas hnRNPs antagonize an interaction of the spliceosome with the pre-mRNA, resulting in exon skipping.19 An increasing number of recent examples showed that in addition to proteins, noncoding RNAs can bind to regulatory cis-elements and change splice site selection.20 The presence of splicing enhancers can explain the action of synonymous mutations in many cases, if a mutation does not change the predicted open reading frame but causes a disease. In many cases, these mutations act through a change in affinity of a splicing regulatory protein, which causes a modification of alternative exon usage.21 Similar to the splice sites, splicing enhancers and silencers follow only degenerate consensus sequences. In the case of exonic splicing regulatory elements, this allows for the coexistence of splicing signals and code requirements.22

5.10.1.3 Circular RNAs

In nearly all protein-coding mRNAs described so far, the alignment of the exons in mRNAs corresponds to the corresponding sequence in the DNA. However, more than 25 years ago, mRNAs with “scrambled exons” were reported, that is, rather than having exon 1-2-3, mRNAs with exons 1-3-2 were described.23 This result was not investigated further, since the genesis of these exons had not been elucidated at the time. Subsequent advances in RNAseq technology showed that these “scrambled exons” were part of circular RNAs (circRNAs). Circular RNAs are generated from splicing in large lariats or are generated by intramolecular splicing of RNAs that are constrained through base pairing; for example some are due to strong secondary structures that are present in inverted repeat sequences. The formation of circular RNAs due to strong secondary structures might be important for humans, since 11% of the human genome consists of Alu elements.24 Because Alu elements originate in structural 7SL RNA, Alu elements can dimerize through the Alu domains that show self-complementarity. Since Alu elements are primate specific, they may generate primate-specific circular RNAs.
Most circRNAs are less than 1% of the linear mRNAs made from the same locus, but in some cases the circRNA is 10-fold more
abundant than the linear mRNA from the locus. Studies in cancer showed that circRNAs change their expression more strongly
than mRNAs upon cell transformation. circRNAs are mainly cytosolic and can be translated into protein. In general, circular
RNAs are much more stable than linear mRNAs, as they will not be degraded by the exosome, which acts like an exonuclease.
Due to their circular structure that recapitulates the binding of the poly(A) tail to the mRNA cap in linear RNAs, circRNAs are potentially
excellent ribosomal substrates, allowing for efficient translation in vitro and in vivo. Thus, in addition to AS, the generation of
round RNAs, performed by the spliceosome, greatly increases the coding capacity of the human genome.

5.10.1.4 The Spliceosome

As discussed above the splicing reaction of the major spliceosome is performed by the spliceosome, an RNA–protein machine of
about 4.8 MDa. The spliceosome contains at least 170 proteins that associate with five RNAs (U1, U2, U4, U5, and U6), generating
the U1, U2, U4, U5, and U6 snRNPs. Similar to the ribosome, these snRNPs assemble in a step-wise fashion on the pre-mRNA,
and rearrangements of the snRNPs bring together the splice sites allowing for the catalysis to be performed. The substrate of the spliceosome is pre-mRNA that is bound to many regulatory factors, such as SR-proteins and hnRNPs, which expose specific splice sites to the spliceosome. In most human internal exons, the 3' and 5' splice sites are paired across the exon through an interaction of U1 and U2 snRNPs, which is referred to as exon definition. In yeast and in a subset of large human exons, the splice site pairing occurs across introns, which is referred to as intron definition. Most of our understanding of the splicing reaction comes from in vitro splicing assays that employ an in vitro RNA with two exons having one splice site each flanking a short intron. These in vitro substrates are recognized according to the intron definition model.

In spliceosome assembly, first U1 snRNP binds to the 5' splice site, aided by the base complementarity between the U1 snRNA and the 5' splice site (Fig. 2). The branch-point-binding protein (BBP or SF1) binds to the branch point, aided by a protein interaction of RNA recognition motif 3 (RRM3) of the U2 auxiliary factor (U2AF65) binding to the polypyrimidine tract of the 3' splice site. This protein complex is called the E (early) complex. In the next step the U2 snRNP associates with the complex and replaces BBP, forming the prespliceosome "complex A." The RNA component of U2 snRNP binds to the branch-site, which causes the branch-adenosine to bulge out of a short helix. A protein interaction between the U2 component SF3B1 (SAP155) and the RRM3 of U2AF helps to displace the BBP from its bound form with U2AF65. The entry of the U6/U4/U5 tri-snRNP, catalyzed by prp28, forms the precatalytic spliceosome "complex B." Several RNA helicases (brr2, snu114, and prp2) cause a rearrangement in this complex, which releases U4 and U1, and the U6 snRNA binding to the 5' splice site. In this activated spliceosome (B* complex), the branch point is brought into close proximity to the 5' splice site, allowing for the formation of the lariat in the first step of splicing, leading to the catalytic "C complex." The C complex contains the free first exon, and the lariat bound to the second exon. Using the prp8, prp16, prp18, and slu7 components, the second step of splicing, which involves the ligation of the two exons, is carried out. Another RNA helicase, prp22 releases the remaining U6/U2/U5 snRNPs and the lariat from the spliced RNA in the postspliceosome complex. After catalysis, the U6, U2, and U5 snRNPs are removed from the lariat using snu114, Brr2, and prp43, are then recycled for another round of catalysis, and the intron is degraded. In vitro studies using small synthetic U2 and U6 snRNA fractions can generate a phosphodiester bond using a lariat-analog as a substrate, indicating that the spliceosome is a ribozyme.

5.10.1.5 The Spliceosome is Linked to Other Nuclear Processes

The mechanism of the spliceosome has been determined biochemically using a nuclear extract made under high salt conditions. When spliceosomes are prepared using physiological salt conditions larger spliceosome complexes, termed supraspliceosomes are observed. These preparations contain other mRNA processing factors, showing that pre-mRNA splicing in vivo is linked to other RNA processing steps. Pre-mRNA splicing occurs during RNA pol II transcription, and there is evidence for cotranscriptional splicing. For example, a fast-moving RNA polymerase promotes alternative exon skipping and snRNPs and processing factors are loaded on the carboxy terminal domain of RNA pol II prior to transcription. However, since RNA pol II has an elongation rate of 3–4 kb/min, large introns are likely spliced out after transcription. There is accumulating evidence that chromatin structure is linked to exon selection. The average length of a human cassette exon of 145 nt corresponds well with the length of DNA in a nucleosome and in fact, splice sites preferably correspond to DNA located at the end of a nucleosome. Histone modifications, such as H3K4me3, assist in the recruitment of U2 snRNP components to sites of active transcription, which could promote exon recognition. The U2 snRNP component SF3B1 physically interacts with histone H3, suggesting a physiological link. It is possible that pre-mRNA splicing affects chromatin changes, in return, as DNA encoding exons are characterized by the H3K36me3 modification. This modification is lower in DNA corresponding to alternative exons, when compared to constitutive exon. Since on average alternative exons of the same pre-mRNA assemble less spliceosomes than the constitutive exons, this suggests that the activity of the spliceosome is reflected in the H3K36me3 modification. Components of the spliceosome also participate in transcriptome surveillance. The U1 snRNA is more abundant than the other spliceosomal U2, U4, U5, and U6 snRNAs. U1 snRNA binds to specific sites on pre-mRNA that are not recognized as splice sites and which prevents premature cleavage and polyadenylation.

5.10.1.6 Mutations in Spliceosomal Proteins

Similar to the ribosome, the spliceosome is evolutionary, highly conserved from yeast to humans. Underlining its central role in metazoans germ line mutations in core spliceosomal proteins are rare and include PRP3, PRP6, PRP8, Brr2, and PRP31, causing dominant nonsyndromic retinitis Pigmentosa. Mutations in the minor spliceosome small nuclear RNA U4atac cause microcephalic osteodysplastic primordial dwarfish type I (Table 1). In contrast to germ line mutations, numerous somatic mutations of core spliceosomal and pre-mRNA-associated proteins have been detected in cancer cells using RNAseq (Table 2). Interestingly, they cluster in components that recognize the 3' splice site, namely, U2AF35 (U2AF1), ZRSR2, a binding partner for U2AF, and the U2 snRNP component SF3B1 (Fig. 4). Similarly, mutations in several, but not all proteins binding to splicing regulatory elements on the pre-mRNA have been detected. They include SRSF1 (SF2/ASF); SRSF2 (SC35). These proteins generally promote exon recognition. In addition, SRSF1 plays a role in genome stability.

5.10.1.7 Mutations in Pre-mRNA

In addition to mutations in the spliceosome, mutations on the pre-mRNA that affect splicing are associated with cancer (reviewed in Refs. 61 and 62). Lists of alternative exons deregulated in different cancer cells have been compiled in the literature
About 20% of point mutations in cancer are synonymous mutations, that is, mutations located in protein coding exons that do not affect the reading frame. Up to 8% of these somatic mutations are selected in cancer cells, suggesting that they are functional. They frequently affect splicing regulatory sequences (exonic enhancers or silencers) and thus contribute to the deregulation of splicing. Examples include BRCA2 in breast cancer, APC in familial adenomatous polyposis, and TP53 in multiple tumors. A much larger number of point mutations are found in intronic regions, but they are mechanistically poorly understood. The best understood effects are the alterations of splice sites. Mutations affecting the conserved AG-GT splice site dinucleotides always abolish usage of the affected splice site, which usually results in the skipping of the exon unless a cryptic splice site is activated. Mutations outside the conserved AG-GT dinucleotides mostly modulate exon usage. However, the mechanistically less well understood mutations are associated with deep intronic alterations that can also affect splice site usage in cancer, usually due to activation of cryptic exons, for example in BRCA2.

5.10.2 The Spliceosome as a Cancer Target

Aberrations in pre-mRNA splicing, in the case of MDM2, for example, were recognized as a feature of tumors as early as 1996, and reviewed in 2002. Subsequently, it was recognized that aberrant pre-mRNA splicing was a global property of the tumor transcriptome; these observations were reviewed as early as 2004 with the prediction that this understanding might lead to new opportunities for the diagnosis and treatment of cancer. However, despite much suggestive data, it was difficult to prove that splicing aberrations are actually oncogenic. In the past few years, much more information has been obtained through genomic and transcriptomic studies of tumors. Such splicing aberrations are now recognized as another predominant feature of cancer and the spliceosome is now seen as a validated antitumor target. In the following, we will explore the history and the evolution of thought regarding the spliceosome as it is relevant to cancer medicinal chemistry by the exploration of applied and basic research studies, which hold the potential to initiate some transformative changes in the treatment of certain cancers.

In parallel to the work described above, strong evidence has continued to mount that aberrant splicing of pre-mRNA is a driver of tumorigenesis and that modulation of this process may be a valid target for cancer therapy. It has been known for some time that the deregulation of exon usage is a hallmark of cancer and recent groundbreaking discoveries have identified recurrent mutations in SF3B1 and/or other splicing factors in multiple forms of cancer including: myelodysplastic syndromes (MDS), chronic lymphocytic leukemia (CLL), acute myeloid leukemia (AML), breast cancer, lung adenocarcinoma, and uveal melanoma. These genetic studies have also fueled complementary research in the therapeutic significance of spliceosome recurrent mutations. Very recently, the selective sensitivity of tumors to agents that target SF3B has also been linked to overexpression of MYC and published as part of a collaborative multidisciplinary effort.

### Table 1: Germ-line mutations of the spliceosome leading to human disease

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRP3</td>
<td>Part of the U4/U5/U6 tri-snRNP</td>
<td>Retinitis pigmentosa</td>
</tr>
<tr>
<td>U4/U6 SnRNP 90 KDa Protein</td>
<td>Part of the U4/U5/U6 tri-snRNP</td>
<td>Retinitis pigmentosa</td>
</tr>
<tr>
<td>PRP6</td>
<td>Part of the U4/U5/U6 tri-snRNP</td>
<td>Retinitis pigmentosa</td>
</tr>
<tr>
<td>PRP8</td>
<td>Scaffold for snRNPs during the splicing reaction</td>
<td>Retinitis pigmentosa</td>
</tr>
<tr>
<td>US SnRNP-Specific Protein (220 KD)</td>
<td>RNA helicase</td>
<td>Retinitis pigmentosa</td>
</tr>
<tr>
<td>BR2</td>
<td>US SnRNP-Specific 200 KDa Protein</td>
<td>Retinitis pigmentosa</td>
</tr>
<tr>
<td>PRP31</td>
<td>U4/U6 SnRNP 61 KDa protein</td>
<td>Retinitis pigmentosa</td>
</tr>
<tr>
<td>U4atac</td>
<td>snRNA in the minor spliceosome</td>
<td>MOPD type 1</td>
</tr>
</tbody>
</table>

MOPD, microcephalic osteodysplastic primordial dwarfism

### Table 2: Somatic mutations of the spliceosome in cancer cells

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF3B1</td>
<td>U2 component, recognition of 3’ splice site</td>
<td>Chronic lymphocytic leukemia, myelodysplasia</td>
</tr>
<tr>
<td>SF3A1</td>
<td>U2 component, recognition of 3’ splice site</td>
<td>Myelodysplasia</td>
</tr>
<tr>
<td>PRPF40B</td>
<td>U1 and U2 component</td>
<td>Myelodysplasia</td>
</tr>
<tr>
<td>SF1</td>
<td>Recognizes branchpoint, 3’ splice site</td>
<td>Myelodysplasia</td>
</tr>
<tr>
<td>U2AF35</td>
<td>recognition of 3’ splice site</td>
<td>Myelodysplasia</td>
</tr>
<tr>
<td>U2AF65</td>
<td>recognition of 3’ splice site</td>
<td>Myelodysplasia</td>
</tr>
<tr>
<td>ZRSR2</td>
<td>Second step of splicing in major spliceosome; recognition</td>
<td>Myelodysplasia</td>
</tr>
<tr>
<td>SRSF2 (SC35)</td>
<td>RNA-binding motif and serine/arginine rich 2</td>
<td>Myelodysplastic syndromes</td>
</tr>
</tbody>
</table>

Targeting Pre-mRNA Processing in Cancer
5.10.2.1 Mechanism of the Selective Antitumor Activity of Spliceosome Targeted Agents

5.10.2.1.1 Synthetic lethal spliceosome pairs

Since the discovery of the first small molecule antitumor agents that target the spliceosome, there has been substantial confusion in regard to the mechanism of selective activity of the spliceosome-targeted agents. This confusion has been in part due to numerous early reports describing these agents as "splicing inhibitors," which led in turn to the development of an incorrect conceptual framework for understanding the true mechanism of action of these agents. Though it is now well established that these agents act by the modulation of alternate splicing due to a reduction in the fidelity of 3'-splice site selection this confusion remains prevalent even in some recent literature reports. While many mechanisms may play a role in the selective killing of tumor cells by the agents discussed below, it is becoming clear that nononcogene addiction is an important mechanism that can explain the selectivity in many cases.86

The nononcogene addiction concept provides a platform for the exploration of a large number of potential targets for anticancer therapies, whose change of function could selectively reverse oncogenic phenotypes. Several groups have carried out synthetic lethal RNA interference screens designed to identify genes whose silencing is specifically cytotoxic for cells carrying particular oncogenic mutations, in order to take advantage of this concept in the development of new anticancer approaches and in the elucidation of the mechanism of action of antitumor agents, including those agents that modulate splicing. These latter studies have relied on the use of "synthetic lethal" genetic studies using various screening strategies (e.g., siRNA and shRNA libraries that specifically "knock down" genes) to elucidate the aberrant splicing dependence of tumors bearing certain mutations. This topic has very recently been authoritatively reviewed.87

Fig. 4  Recurrent mutations in the drug protein target SF3B1. The scheme depicts the domain organization of splicing factor 3B subunit 1 (SF3B1) according to Ref. 52. The amino-terminal region is involved in interactions with other splicing factors, whereas the carboxy-terminal domain encompasses 22 HEAT (huntingtin, elongation factor 3, protein phosphatase 2A, target of rapamycin 1) repeats. Repeats harboring mutations found in chronic lymphocytic leukemia and/or myelodysplastic syndrome tumor samples are represented in orange, and the most frequent mutation and its corresponding repeat are represented in red.46,53 U2AF65, U2 snRNP auxiliary factor of 65 kDa.4 Reprinted by Bonnal, S.; Vigevani, L.; Valcarcel, J. The Spliceosome as a Target of Novel Antitumour Drugs. Nat. Rev. Drug Discov. 2012, 11, 847–859, with permission from Macmillan Publishers Ltd.
One of the first reports of the synthetic lethal dependence of a specific mutation on splicing function is that of Hubert et al.\textsuperscript{85} These investigators searched for glioblastoma multiforme (GBM)-lethal genes using genome-wide RNAi screens in multiple GMB stem cells (GSCs), patient isolates, and neural stem cells (NSCs) in order to classify the genes that are differentially required for GSC expansion. These studies revealed that the plant homeodomain-finger domain protein (PHF5A) was differentially required for expansion and viability of several GSC isolates. Further, molecular studies demonstrated that GSCs have a novel requirement for PHF5A activity. Knockdown of PHF5A resulted in splicing changes in thousands of essential genes, some of which are expected to alter cell division and growth. Also, these authors showed that sensitivity to U2snRNP perturbation could be established in NSCs and fibroblasts overexpressing MYC. Although MYC is not frequency amplified in GBM, MYC is coordinately activated by mutations in p53 and PTEN. In summary, this study established that patient-derived GSCs are vulnerable to perturbation in splicing, which results in a reduction in GSC viability and loss of GBM tumor maintenance. The authors also showed that splicing modulators such as sudemycin C1 showed selective toxicity to GBMs and fibroblasts overexpressing MYC. Although standard of care therapies is ineffective against GBM, the authors concluded that targeting PHF5A and/or U2 snRNP activity offers a new therapeutic inroad for this GBM.

Another more recent study also identified sudemycin sensitivity and a synthetic lethal partner in MYC-driven triple negative breast cancers.\textsuperscript{88} In this work, it was found that the spliceosome is a target in MYC-driven cancers. These researchers identified the BUD31 protein as an MYC-synthetic lethal gene in human mammary epithelial cells, and showed that BUD31 was required for spliceosome assembly and catalytic activity. Activity of SF3B1 and other splicing factors associate with BUD31 are also required by oncogenic MYC. MYC activation induces an increase in total pre-mRNA synthesis, suggesting an increased burden on the core spliceosome to process pre-mRNA. In contrast to normal cells, reduction of functional SF3B1 in MYC-hyperactivated cells leads to splicing aberrations, widespread effects on pre-mRNA maturation, and deregulation of many critical cell processes. Genetic or pharmacological modulation of the spliceosome in vivo impairs survival, tumorigenicity, and metastatic inclination of MYC-dependent breast cancers. Importantly, this work showed that a reduction in SF3B1 activity is effective in human MYC-dependent breast cancer patient-derived xenografts (PDX) in vivo using the SF3B1 targeted agent sudemycin D6 (SD6) that was developed in the Webb laboratory.\textsuperscript{88}

5.10.2.1.2 Tumors bearing spliceosome mutations are vulnerable to splicing modulation
Following the identification of many common spliceosome recurrent mutations in the MDS, CLL, and AML, there was initially confusion regarding the functional effects of these mutations, since little was known regarding the precise functional consequences of these splicing mutations. The observed recurrent mutations are consistently heterozygous and only very rarely occur together, indicating that one wild-type copy of the normally functional gene product is essential. Given that the spliceosome has important regulatory functions, the standard expectations of “gain of function” or “loss of function” are likely undefined with respect to the splice-site regulatory subunits, given the current understanding that the spliceosome components that are mutated are involved in the fidelity of 3′-splice site recognition. Thus, for these proteins a different concept: change of function, must be considered. Many researchers hypothesized that such cancers might be particularly sensitive to agents that target SF3B1 since these mutations seemed to produce an oncogenic splicing program. Thus, agents such as SD6 or E7107 were used to test this hypothesis (see Figs. 5 and 6).

One of the first such studies that was reported found that the spliceosome modulators sudemycin D1 and D6 showed selective cytotoxicity in primary CLL cells when compared with normal lymphocytes and tumor cells from other B-lymphoid cancers, with a slight bias for CLL cases bearing spliceosome mutations.\textsuperscript{107} This work reported that sudemycin exhibited antitumor activity in
NOD/SCID/IL2Rγ−/− (NSG) mice engrafted with primary cells from CLL patients. This study showed a remarkable selectivity of the SF3B1 targeted agent SD6 for most of the CLL genotypes, which may indicate a common mechanism is driving CLL in addition to the spliceosome mutations and that this common mechanism results in selective susceptibility to any additional reduction in 3′-splice site fidelity. In another conceptually related publication, the authors tested this hypothesis using mice genetically engineered to express a mutated allele of SR splicing factor 2 (Srsf2P95H), which is recurrent in individuals with MDS and AML, in an inducible manner in hematopoietic cells. These mice rapidly succumbed to fatal bone marrow failure, demonstrating that Srsf2-mutated cells depend on the wild-type Srsf2 allele for survival. These authors also showed that treatment with the spliceosome modulator E7107 specifically showed anticancer activity in both isogenic mouse leukemias and PDX AMLs carrying spliceosomal mutations. Thus, the authors of this publication concluded that:

Collectively, these data provide genetic and pharmacol evidence that leukemias with spliceosomal gene mutations are preferentially susceptible to additional splicing perturbations in vivo as compared to leukemias without such mutations. Modulation of spliceosome function may thus provide a new therapeutic avenue in genetically defined subsets of individuals with MDS or AML.

5.10.2.2 Natural Products and Derivatives Targeting SF3B1

In 2007, two research groups (from Astellas Pharma, Inc. and Eisai Co, Ltd., respectively) independently reported two structurally dissimilar bacterial natural products, FR901464 (FR) and pladienolide (Fig. 5), both targeting a similar site on the SF3B subunit of the spliceosome. This work proved to be foundational for the field of small molecules that target SF3B1. Subsequent work inspired by these 2007 publications reported additional natural products that were shown to target the SF3B subunit. These subsequent reports include herboxidiene (GEK1A) (isolated from Streptomyces sp. A7847) and the thailanstatins (isolated from Burkholderia thailandensis). A similar interaction with the SF3B subunit (which is part of the spliceosomal U2 snRNP that recognizes the branch point and the 3′-splice site) is also likely for the macrolide natural product FD-895, given its potent biological activity.

Fig. 6 Active totally synthetic analogs of splicing modulatory natural products discussed in the text. The left panel shows the Burkart analogs, the weakly active Webb pladienolide analog, and 6-norherboxidiene, meayamycin, and the sudemycins.
and its high structural and pharmacophore similarity to pladienolide. Many of these bacterial fermentation products show cytotoxic IC50S in the low nanomolar range in several tumor cell lines and have been reported to have a similar distinctive effect on the cell cycle in mammalian cell lines, including cell cycle arrest in the G1 and G2/M phases. Several of these natural products have also been reported to show potent antitumor activity in vitro and in vivo.

Work in this area led to the development of the semisynthetic pladienolide analog E7107 (Fig. 5) that entered Phase I clinical studies, without the benefit of many subsequent recent discoveries relevant to mechanism of action, tumor selectivity, and patient stratification. The clinical trials of E7107 were suspended and never reinitiated due to the ocular toxicity of this compound. These results strongly supports the use of pharmacodynamic assays that have recently been reported, as an integral part of the splicing modulatory drug development process, as was used for SD6, since the pharmacodynamics for this class of drugs has unique features. The most notable feature of this class of drugs is the long-lasting cytotoxic effects following drug "washout" in vitro. Though this washout effect is most notable for tumors, we inferred that this is likely to also be the case for sensitive, rapidly dividing normal tissue. We concluded that pulsatile treatment (via slow intravenous infusion) with a fast-acting drug, which possesses a relatively short half-life, is a critical criterion to use in the development of drugs in this class. In contrast to this requirement, observed in our studies, E7107 was optimized on the basis of developing a longer half-life when compared to the parent natural product (pladienolide) series.

Though researchers initially characterized these compounds as “splicing inhibitors,” the mechanism was subsequently elucidated and these compounds are now understood to effect alternate splicing through altering the 3'-splice site selection fidelity of the spliceosome and are correctly understood to be splicing modulators and not general inhibitors of splicing. One of our important contributions to this field has been our publication of cell-based splicing assays that use a PCR readout, or a luciferase reporter, and the observation that the cytotoxicity potency of a range of natural or synthetic splicing modulators correlates to the potency for splicing modulation in this assay. It has also been shown that the potencies of pladienolide, herboxidiene, and sudemycin in cell-based cytotoxicity assays correlate with the potencies in a cell-based luciferase splicing reporter assay. Evidence has been presented that this latter cell-based splicing reporter assay specifically reports on-target spliceosome modulatory pharmacodynamic activity in real time. The structures for the compounds discussed above are presented in Fig. 5.

5.10.2.3 Synthetic natural product analogs targeting SF3B1

The initial analogs of the natural product FR901464 were prepared as part of work that examined novel total synthetic routes to this compound first by the Jacobsen group, and then by the Koide group who also developed more chemically stable analogs (meayamycins). and explored the structure-activity for a set of synthetic analogs of this compound class. FD-895 is another natural product that is structurally related to pladienolide, which shows similar splicing modulatory and antitumor activity, and was the target of the successful total synthesis and analoging effort by the Burkart group. Most remarkably, the Burkart group recently developed a completely novel, active structural analog, which replaces the macrolide component of FD-895 with a sugar scaffold. This latter work is a tour de force and truly an extraordinary example of natural product molecular design.

As part of an effort to develop a class of drug-like synthetic spliceosome modulators the Webb group reported the design and synthesis of FR analogs that contain only 3 chiral centers (the sudemycins), and several herboxidiene analogs (including a pladienolide–herboxidiene hybrid), which included many new active compounds that effectively modulate alternate splicing. Additionally, we recently reported our results from a genome-wide array analysis of sudemycin treated tumor cells; we found that sudemycins cause a rapid wide-ranging change in alternate pre-mRNA splicing. This same paper also showed that a biotin-labeled sudemycin probe directly interacts with the SF3B1 protein. Our lead optimization work has primarily focused on the sudemycins, which are chemically stable synthetic molecules that were designed based on the known FR structure–activity relationships and the application of a hypothetical consensus pharmacophore model that is derived from molecular overlays of FR and PD. This work ultimately led to SD6, which is currently in preclinical Investigational New Drug (IND) development, with clinical studies for the treatment of myelodysplastic syndromes expected to start in 2018. The structures for the compounds discussed above are presented in Fig. 6.

5.10.2.4 Inhibitors of the CDC-Like Kinases

Compounds that inhibit members of the cdc-like kinase (CLK) family of kinases are known to modulate pre-mRNA splicing through the inhibition of SR protein phosphorylation. The medicinal chemistry of this emerging area has been extensively reviewed. Several diverse marine natural products and synthetic compounds have been found to be inhibitors of the CLK family (Figs. 7–9). Several groups have shown that selective CLK inhibitors modulate splicing through their action on spliceosome components. This area is nascent and may be fertile ground for the discovery of new compounds for cancer and other diseases that are associated with splicing aberrations. With the recent insights on the biology of the spliceosome medicinal chemistry in this area, as discussed above, drug discovery targeting the various CLKs (especially CLK1 and CLK2) is likely to expand in near future.
Fig. 7  Examples of marine natural products targeting CLks.122

Fig. 8  A representation of the structure of debromohymenialdisine (K0010) bound to CLK1 using the published coordinates,125 highlighting the hydrogen bonds to the ligand (as blue lines). This figure was prepared using UCSF Chimera.126

Fig. 9  Synthetic CLK, DYRK1A, and SRPK1 inhibitors showing cell-based activity121: KH-CB19,124 Araki Cpd-2,123 INDY,127 and SRPIN340128
5.10.3 Conclusions

As discussed earlier, combined progress in natural product screening, target identification, spliceosome-related medicinal chemistry, and high-throughput transcriptome sequencing has led to a remarkable convergence of independent research areas, which have simultaneously identified new oncology drug targets and new small-molecule therapeutic agents. There is good reason to believe that these insights will lead to an expansion of spliceosome medicinal chemistry and the development of many new selective and highly effective chemotherapeutics for a range of cancers.

References


