Acquired Mutations That Affect Pre-mRNA Splicing in Hematologic Malignancies and Solid Tumors

Linda M. Scott and Vivienne I. Rebel

The application of next-generation sequencing technologies to interrogate the genome of human hematologic malignancies is providing promising insights into their molecular etiology and into the pathogenesis of seemingly unrelated malignancies. Among the somatic mutations identified by this approach are ones that target components of the spliceosome, a ribonucleoprotein complex responsible for the posttranscriptional processing of primary transcripts to form mature messenger RNA species. These mutations were initially detected in patients with chronic lymphocytic leukemia or a myelodysplastic syndrome, but can also occur at relatively high frequency in some solid tumors, including uveal malignant melanoma, adenocarcinoma of the lung, and estrogen receptor–positive breast cancers. Their presence in a variety of malignancies suggests that the spliceosomal mutations may play a fundamental role in defining the malignant phenotype. The development and testing of drugs that eliminate cells bearing a spliceosomal mutation, or normalize their altered transcript splicing patterns, are therefore a priority. Here, we summarize the effects of spliceosome-associated mutations on transcript processing in vitro and in vivo, and their impact on disease initiation and/or progression and patient outcome. Moreover, we discuss the therapeutic potential of compounds already known to target splicing factor 3B subunit 1 (SF3B1), an essential component of the spliceosome that is frequently mutated.

The Process of mRNA Splicing

Posttranscriptional processing of primary transcripts to form mature mRNA species involves the removal of intronic sequences by the spliceosome, a complex that consists of five small nuclear ribonucleoproteins (snRNPs) and between 100 and 300 associated proteins. The “major” or U2-type spliceosome, which consists of the U1, U2, U4, U5, and U6 snRNPs, catalyzes the vast majority of transcript splicing events, whereas the “minor” or U12-type spliceosome utilizes the U12 snRNP to mediate splicing of approximately 800 specific transcripts. The complex multistep process that enables splicing has been comprehensively described in two excellent review articles (8,9). For the purposes of this article, however, we briefly summarize the steps involved in “major” transcript splicing. As shown in Figure 1, this process is initiated by the spliceosome assembly phase, which includes the sequential formation of three complexes: an early or “E” complex, an “A” complex, and a “B” complex. “E” complex formation results from the binding of a U1 snRNP to the 5′ splice site, and the U2-associated factor, U2AF, to the 3′ splice site and adjacent polypyrimidine tract. The U2AF protein is a heterodimer consisting of a U2AF1 subunit (also called
Figure 1. Stepwise removal of intronic sequences from a transcript by the spliceosome. Posttranscriptional processing of primary transcripts to form mature mRNA species involves the removal of intronic sequences by the spliceosome, a complex generally consisting of five small nuclear ribonucleoproteins (snRNPs; U1, U2, U4, U5 and U6) and a large number of associated proteins. The splicing process is initiated by the binding of U1 to the 5′ splice site, which includes a GU dinucleotide sequence, and weak binding of the U2-associated factor (U2AF), to the 3′ splice site (AG) and adjacent polypyrimidine (pY) tract, resulting in the formation of an E complex. In the next phase, an ATP-dependent alteration in the transcript positioning brings into close proximity the two exons (labeled here as I and II, respectively), facilitating binding of the U2 snRNP to the branchpoint sequence (denoted by #) to generate an A complex. The B complex is formed upon the addition of a U4/U5/U6 snRNP triad to the spliceosome, which then undergoes rearrangement to form the catalytically active C complex, with loss of U2AF and the U1 and U4 snRNPs. Finally, the intervening intron is removed via two transesterification steps, so that exons I and II lie immediately adjacent to one another in the mature transcript.

U2AF35) and a U2AF2 (or U2AF65) subunit, which mediate binding to the 3′ splice site and the polypyrimidine tract, respectively. An ATP-dependent alteration in transcript positioning then brings together the proximal and distal exons, facilitating the binding of the U2 snRNP and generation of an “A” complex. This snRNP consists of a 12S RNA subunit and the SF3A and SF3B multiprotein complexes, with SF3B1 mediating binding to the intronic branchpoint sequence. This motif serves as a docking site for the spliceosomal subunits essential for the transesterification steps involved in splicing. The “B” complex is then formed by the addition of the U4/U5/U6 snRNP triad, which enables spliceosome activation, in which an ATP-dependent conformational rearrangement forms the catalytically active “C” complex, with release of U2AF and the U1 and U4 snRNPs. Two transesterification steps result in excision of the intron, and ligation of the exons such that the proximal and distal exons lie adjacent to one another in the mature transcript.

Most, if not all, of the spliceosomal proteins involved in the removal of introns from the primary transcript also influence patterns of alternate splicing. Members of the serine/arginine (SR) family play a particularly crucial role in the regulation of this process (10). The SR proteins contain one or more RNA binding domains (RBDs) that mediate their ability to interact with intronic and exonic splicing enhancers (ISEs and ESEs, respectively) present in many primary transcripts. Binding of SR proteins at these enhancer elements enables recruitment of the U1 snRNP and U2AF to the neighboring 5′ and 3′ splice sites, respectively, promoting the inclusion of that exon into the mature transcript. However, this process is regulated by posttranslational modification; the methylation, acetylation, or phosphorylation of SR proteins alters their ability to bind ISEs and ESEs. Heterogeneous nuclear ribonucleoprotein complexes instead bind to intronic or exonic splicing silencer elements (ISSs and ESSs), blocking the recruitment of the U1 snRNP and U2AF by a mechanism that has not been fully defined. Most primary transcripts that have been studied contain multiple ISE/ESE or ISS/ESS elements; it is the overall balance between the activities at these two antagonistic classes of regulatory element that determines whether or not a particular exon will be included in the mature transcript.

Defects in the Constituents of the Spliceosome in Myeloid Malignancies

The first example of somatic mutations affecting components on the spliceosome was reported in 2011, when whole exome sequence analysis was performed on paired tumor and normal samples from 29 patients with MDS (5). Among this cohort, 16 individuals had an acquired mutation that affected a component of the splicing machinery (namely PRPF40B, SRSF2, SF3A1, SF3B1, U2AF1, or ZRSR2). The mutation status of these six proteins, and another three spliceosome-associated factors (SF1, SRSF1, and U2AF2), was evaluated in an additional 582 samples from patients with a myeloid malignancy. Mutually exclusive mutations were detected in 209 cases, including in 6.6% of patients with de novo AML and...
9.4% of patients with a myeloproliferative neoplasm (MPN), and occurred most frequently in the subset of MDS patients that are positive for ringed sideroblasts, which are morphologically distinct erythroid precursor cells in which abnormal crystalline iron deposits surround the mitochondria. This included 83% of patients with refractory anemia with ringed sideroblasts (RARS), and 76% of those with refractory cytopenia with multilineage dysplasia and ringed sideroblasts (RCMD-RS). SRSF2 mutations were also detected in more than a quarter of patients with chronic myelomonocytic leukemia, a clonal hematopoietic stem cell (HSC) disorder with characteristics of both MDS and an MPN. Subsequent studies (6,7,11–15), several of which are summarized in Table 1, have reported comparable incidences of spliceosome mutations in adult patients with a myeloid malignancy. One of these studies also showed that mutations may affect the LUC7L2, PRPF8, SF3B3, SON, or SRSF6 spliceosome proteins (11). In contrast, spliceosomal mutations are relatively rare in cases of pediatric MDS (occurring in 1 of 255 cases) (16), suggesting that the molecular pathogenesis of juvenile and adult MDS are distinct.

The targets for the majority of spliceosome-associated mutations in MDS are components of the “E” and “A” splicing complexes (Figure 1). U2AF, which is a heterodimer of the U2AF1 and U2AF2 proteins, binds to the 3′ splice site, bringing together adjacent exons and facilitating the binding of the U2 snRNP. This snRNP includes the SF3A and SF3B complexes; the former contains three protein subunits (SF3A1–3), and the latter consists of five (SF3B1–5). ZRSR2 associates with the U2AF1 protein and plays a role in the recognition of the 3′ splice site, whereas SF1 recognizes the branchpoint sequence and is required for the generation of an “A” complex. The SR proteins, including SON, SRSF1, SRSF2, and SRSF6, interact with other constituents of the spliceosome via their arginine/serine-rich (RS) domains, forming a bridge that links together the 5′ and 3′ splice sites. LUC7L2, PRPF8, and PRPF40 are, however, constituents of snRNPs other than U2. Relatively little is known about the function of LUC7L2 or PRPF40B, which have homology to the yeast LUC7 and PRP40 splicing factors, respectively, although LUC7L2 has a C2H2-type zinc finger and an RS domain that suggest it is involved in mediating interactions between spliceosome components. Both proteins are present in the U1 snRNP, whereas PRPF8 is a component of the U5 snRNP, where it interacts with the 5′ and 3′ splice sites to align the proximal and distal exons for subsequent ligation. PRPF8 is therefore essential for the second transesterification step in pre-mRNA splicing.

The spliceosome mutations identified are predominantly simple amino acid substitutions, although small in-frame deletions within SRSF2 have been noted in patients with myelofibrosis or de novo MDS (17,18). The location of the most frequently occurring alterations is shown in Figure 2. Most disease-associated mutations can be localized to one of the functionally important domains in SF3B1, U2AF1, or SRSF2. In SF3B1, most mutations map to one of the 22 HEAT Huntingtin, Elongation factor 3, protein phosphatase 2A, Targets of rapamycin 1 repeats present in this polypeptide, and which mediate its interaction with other proteins. They cluster around residues 625, 666, and 700, with a K700E substitution occurring most frequently. Mutations also commonly target either of the zinc fingers present in U2AF1, with amino acid substitutions occurring at residues S34 and Q157. In direct contrast, mutations do not affect the RBD or zinc finger of SRSF2, but instead target residue P95, which lies between these domains. There are no mutational hotspots in ZRSR2, with alterations occurring throughout the protein; several are amino acid substitutions that affect the RBD or distal zinc finger motifs, although many are nonsense or frameshift mutations that result in protein truncation.

**SF3B1 Mutations and Patient Outcome in Patients Diagnosed With CLL**

Acquired SF3B1 mutations also are present in a substantial proportion of patients with CLL (19–23) (Table 2). In contrast, the majority of the other spliceosome mutations associated with a myeloid malignancy are conspicuously absent, although SRSF1, SRSF7, or U2AF2 mutations can occur on occasion (22,23). Rare CLL-associated mutations may also target spliceosomal components that are not affected in patients with MDS, including two that bind to members of the SR family: SRPK1 and SRRM1 (23).

The CLL-associated SF3B1 mutations all localize to the HEAT repeats, occurring in the same regions as the mutations present in patients with MDS (Figure 2). Their frequency ranges substantially between different cohorts (Table 2), which likely reflects both the genetic diversity of patients within each cohort and their exposure or response to chemotherapy. For example, the frequency of SF3B1 mutations in fludarabine-refractory patients was three times greater than that of untreated patients (20). This difference is not solely attributable to a more advanced disease stage in the former grouping, as patients who had undergone disease transformation to Richter syndrome had a mutation frequency comparable to that of newly diagnosed patients. SF3B1 mutations are more prevalent in patients with a deletion of chromosome 11q22 or mutation in the *ATM* locus, which lies within this region (22), and in patients without an *IGHV* rearrangement (23). Accordingly, mutation-positive patients tend to be diagnosed with a more advanced disease than those lacking an SF3B1 mutation, and have a statistically significant lower 10-year overall survival rate (23–25). However, the presence of an SF3B1 mutation did not appreciably alter overall survival or event-free survival rates in high-risk CLL patients receiving an allogeneic hematopoietic stem cell transplant (21). Analysis of the transcriptome of patients with or without an SF3B1 mutation showed that the former express several alternatively spliced and prematurely truncated transcripts that are not present in the latter grouping (23). Among the most clinically relevant targets identified by this approach was FOX1, a transcription factor associated with the pathogenesis of diffuse large B-cell lymphoma (26). A unique FOX1 transcript that is specific to CLL patients who have an SF3B1 mutation encodes a shortened protein that lacks several of the PEST motifs that mediate its degradation (Figure 3A). This presumably alters the FOX1 expression level in SF3B1 mutation–positive cells, thereby promoting proliferation, although this possibility remains to be tested.

**Spliceosomal Mutations in Nonhematologic Malignancies**

Although the spliceosomal mutations were originally associated with hematologic malignancies, recent next-generation sequencing studies have demonstrated that they may also occur in several
solid tumor types, including neuroblastomas that arise following chromothripsis (27), pancreatic ductal adenocarcinoma (28), and estrogen receptor–positive tumors of the breast (29–31). Most of these affect SF3B1, although mutations within SR proteins (including SRSF1, SRSF12, and SRSF14) have been detected in malignant breast tumors (31). Additional examples of spliceosome-associated mutations occurring in solid tumors are to be expected as the number of whole genome sequences reported in the literature rises. The true relevance of these spliceosome mutations is also only just becoming apparent, with an unbiased analysis of the genomes of 510 breast cancers identifying SF3B1 as one of the 35 most frequently mutated genes in this tumor type (32). Similarly, an independent evaluation of 99 pancreatic tumors identified SF3B1 among the 16 most frequently mutated genes (28).
Acquired heterozygous SF3B1 mutations were recently identified in a sizeable fraction (22 of 105) of patients with uveal malignant melanoma (33). In stark contrast to the mutation distribution seen in patients with a hematologic malignancy, the melanoma-associated mutations target a single codon. This causes the substitution of arginine-625, located in the fifth HEAT repeat of SF3B1, with a cysteine, glycine, histidine, or leucine. Within this cohort, mutations in SF3B1 or BAP1, a nuclear ubiquitin carboxy-terminal hydrolase whose inactivation occurs in the majority of uveal melanoma patients with a high metastatic risk (34), occurred in a mutually exclusive fashion. Accordingly, the SF3B1 mutation-positive patients had a relatively favorable prognosis; compared to mutation-negative patients, they presented at an earlier age, and had tumors with a lower number of undifferentiated epithelial cells (33). SF3B1 mutations do not feature prominently in primary or metastatic melanomas originating in the hair-bearing skin (35), or in patients with cutaneous melanoma (36).

Somatic U2AF1 mutations are also a recurring feature of lung cancer (37), albeit with a lower frequency than the SF3B1 mutations in breast or pancreatic cancers. Overall, 10% of patients with an adenocarcinoma of the lung were positive for a somatic mutation targeting a component of the splicing machinery. Four of the 183 tumor samples interrogated carried the U2AF1S34F mutation originally described in MDS patients. Although the number of mutation-positive cases was small, the presence of a U2AF1S34F mutation was nevertheless associated with a reduction in the progression-free survival rate. An additional patient had a G-to-I substitution at residue 271 of U2AF1, which is located in the RS domain; this variant has not been observed in patients with a hematologic malignancy. In an additional 14 cases, nonsynonymous

Table 2. Incidence of SF3B1 mutations in patients with chronic lymphocytic leukemia*

<table>
<thead>
<tr>
<th>Study</th>
<th>Ref</th>
<th>CLL cohort particulars</th>
<th>Incidence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landau et al., 2013</td>
<td>(19)</td>
<td>160 cases</td>
<td>23 of 160 (14.4)</td>
</tr>
<tr>
<td>Rossi et al., 2011</td>
<td>(20)</td>
<td>301 untreated cases, 59 refractory to fludarabine, and 33 with Richter syndrome</td>
<td>17 of 301 (5.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10 of 59 (16.9)</td>
</tr>
<tr>
<td>Dreger et al., 2013</td>
<td>(21)</td>
<td>100 high-risk cases</td>
<td>2 of 33 (6.1)</td>
</tr>
<tr>
<td>Wang et al., 2011</td>
<td>(22)</td>
<td>91 treated and untreated cases, including 22 with a del(11q)</td>
<td>14 of 91 (15.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8 of 22 (36.4)</td>
</tr>
<tr>
<td>Quesada et al., 2012</td>
<td>(23)</td>
<td>279 untreated cases, including 76 with an IGHV mutation and 73 lacking an IGHV mutation</td>
<td>27 of 279 (9.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 of 76 (7.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>15 of 73 (20.5)</td>
</tr>
<tr>
<td>Oscier et al., 2013</td>
<td>(24)</td>
<td>494 untreated patients recruited to the UK LRF CLL4 trial</td>
<td>73 of 437 (16.7)</td>
</tr>
</tbody>
</table>

* CLL = chronic lymphocytic leukemia; IGHV = immunoglobulin heavy chain variable region.
mutations were present in genes encoding other spliceosomal proteins, including U2AF2 (n = 4), SF3B1 or PRPF40B (n = 3), and PRPF8, SF3A1, or SF3B3 (n = 1).

**Functional Consequences of Spliceosome Mutations In Vitro and In Vivo**

Because of their nature, it was not immediately apparent whether the spliceosome mutations identified in hematopoietic and solid tumors enhanced or impaired pre-mRNA processing, or if they influenced differential splicing patterns. However, the mutually exclusive nature of the spliceosome mutations in patients with MDS suggests that these may have similar effects in vivo, although this has not been determined experimentally. However, insights into the consequence of U2AF1 mutations were provided by experiments testing the in vitro activity of U2AF1 (23). FOXP1w arises from transcript splicing in tumor cells that does not occur in healthy cells: a normally noncoding sequence that includes a termination codon (denoted by an asterisk) is incorporated into the primary transcript instead of exon 20, which encodes the PEST motifs and carboxy terminal domain. The abnormal transcript is not degraded by nonsense-mediated decay; the truncated protein is thought to be stabilized by the absence of PEST motifs, thereby promoting cell proliferation. B) In healthy cells, the overall balance of alternate splicing patterns is governed by the balance between the activities of the splicing enhancer elements (ISE/ESEs) and splicing silencer elements (ISS/ESSs). Examples of normal differential splicing patterns are provided by the MCL-1 locus. The full-length antiapoptotic MCL-1 protein, MCL-1L, is encoded by three exons (labeled here as I, II, and III). A shorter isoform, MCL-1S, arises from a splicing event that excludes exon II from the mature transcript. This alternate splicing replaces the transmembrane domain with a nonsense domain (diamond motif), which renders the protein proapoptotic. The BCL-X gene includes two coding exons (A and B), with the former containing an internal 5′ splice site (ISS). When this is utilized, a shorter mature transcript is produced; this encodes the proapoptotic BCL-XL isoform. Splicing that includes the distal portion of exon A (A′) in the mature transcript instead produces an antiapoptotic BCL-XS isoform. The use of alternative (mutually exclusive) exons may confer distinct properties to the resulting protein: inclusion of exon 9 in the PKM transcript results in expression of pyruvate kinase M1 isoenzyme (PKM1), which promotes oxidative phosphorylation; exon 10 inclusion leads to PKM2 expression, promoting aerobic glycolysis. The balance in alternate splicing in malignant cells may be perturbed by the mutation of components of the spliceosome, favoring the synthesis of isoforms that block apoptosis (MCL-1L or BCL-XL) or provide a metabolic advantage by promoting a high rate of glycolysis (PKM2). In other instances, missplicing may stabilize protein expression levels (FOXP1), or produce truncated transcripts that are degraded via nonsense-mediated mRNA decay. BCL-XL = long isoform of BCL-X (B cell lymphoma-extra; BCL2L1); BCL-XS = short isoform of BCL-X; FOXP1 = forkhead box protein P1; ISS = intronic silencing splicing element; MCL-1 = induced myeloid leukemia cell differentiation protein; PKM-1, -2 = pyruvate kinase isozymes M1 and M2.

Figure 3. Spliceosome mutations may generate protein isoforms that provide cells with a competitive advantage in vivo. A) The mutation of SF3B1 in patients with chronic lymphocytic leukemia leads to the expression of a truncated polypeptide, FOXP1w, which lacks several of the PEST motifs that are present in FOXP1 (23). FOXP1w arises from transcript splicing in tumor cells that does not occur in healthy cells: a normally noncoding sequence that includes a termination codon (denoted by an asterisk) is incorporated into the primary transcript instead of exon 20, which encodes the PEST motifs and carboxy terminal domain. The abnormal transcript is not degraded by nonsense-mediated decay; the truncated protein is thought to be stabilized by the absence of PEST motifs, thereby promoting cell proliferation. B) In healthy cells, the overall balance of alternate splicing patterns is governed by the balance between the activities of the splicing enhancer elements (ISE/ESEs) and splicing silencer elements (ISS/ESSs). Examples of normal differential splicing patterns are provided by the MCL-1 locus. The full-length antiapoptotic MCL-1 protein, MCL-1L, is encoded by three exons (labeled here as I, II, and III). A shorter isoform, MCL-1S, arises from a splicing event that excludes exon II from the mature transcript. This alternate splicing replaces the transmembrane domain with a nonsense domain (diamond motif), which renders the protein proapoptotic. The BCL-X gene includes two coding exons (A and B), with the former containing an internal 5′ splice site (ISS). When this is utilized, a shorter mature transcript is produced; this encodes the proapoptotic BCL-XL isoform. Splicing that includes the distal portion of exon A (A′) in the mature transcript instead produces an antiapoptotic BCL-XS isoform. The use of alternative (mutually exclusive) exons may confer distinct properties to the resulting protein: inclusion of exon 9 in the PKM transcript results in expression of pyruvate kinase M1 isoenzyme (PKM1), which promotes oxidative phosphorylation; exon 10 inclusion leads to PKM2 expression, promoting aerobic glycolysis. The balance in alternate splicing in malignant cells may be perturbed by the mutation of components of the spliceosome, favoring the synthesis of isoforms that block apoptosis (MCL-1L or BCL-XL) or provide a metabolic advantage by promoting a high rate of glycolysis (PKM2). In other instances, missplicing may stabilize protein expression levels (FOXP1), or produce truncated transcripts that are degraded via nonsense-mediated mRNA decay. BCL-XL = long isoform of BCL-X (B cell lymphoma-extra; BCL2L1); BCL-XS = short isoform of BCL-X; FOXP1 = forkhead box protein P1; ISS = intronic silencing splicing element; MCL-1 = induced myeloid leukemia cell differentiation protein; PKM-1, -2 = pyruvate kinase isozymes M1 and M2.
in vivo. Therefore, additional studies are required to fully determine the functional consequences of mutations targeting U2AF1, SF3B1, and other spliceosome-associated proteins.

It is generally accepted that during tumorigenesis, cells of the malignant clone evolve by accruing extra mutations that provide them with a competitive advantage over their sister cells. Surprisingly, then, the presence of a U2AF1 mutation appears to compromise the reconstituting ability of affected HSCs (5). In this study, lethally irradiated mice were transplanted with bone marrow cells transduced with retroviruses encoding wild-type or mutant U2AF1; chimerin in the peripheral blood was assessed over time using the coexpressed green fluorescence protein. Six weeks after transplantation, the proportion of green fluorescence protein–positive cells in the recipients of donor cells expressing U2AF1Q157P, U2AF1Q157R, or U2AF1S34F was less than that of recipients of cells expressing wild-type U2AF1, suggesting that HSC engraftment, proliferation, or differentiation is impaired by mutant U2AF1. However, these observations are in direct contrast to the finding that the acquisition of an SF3B1 mutation in patients with CLL provides mutation-bearing cells with a proliferative advantage (16). A similar analysis in MDS patients may be necessary to determine whether these mutations are advantageous in a myeloid cell context.

Two animal models of SF3B1 haploinsufficiency have provided further insights into the consequences of an acquired SF3B1 mutation in patients with a hematologic malignancy. An N-ethyl-N-nitrosourea–based mutagenesis screen using zebrafish embryos identified a lethal mutant (toastr440) that was characterized by cell death in the dorsal central nervous system and by an absence of circulating blood (39). These animals were found to have a hypomorphic Sf3b1 allele that encodes a truncated, nonfunctional Sf3b1 transcript. The expression level of several critical regulators of neural crest development was dramatically reduced in toastr440 embryos as a result of defective pre-mRNA processing, which in turn caused defects in lineage specification, cell survival, and migration within the neural crest. The blood defect has not yet been assessed, although it is plausible that defects in lineage specification, survival, and/or migration may also be responsible for their hematologic phenotype. Murine gene-targeting experiments revealed that, although Sf3b1-null embryos die at the 32- to 64-cell stage, those heterozygous for an Sf3b1 knockout allele remain viable and healthy (40). As adults, these mice have skeletal anomalies, including a reduced number of ribs, and mild lymphopenia (41). An elevated number of ringed sideroblasts was also noted when Sf3b1-haploinsufficient bone marrow cells were stained with Prussian Blue, a dye that detects the crystalline iron deposits present within this cell type (41). Treatment of healthy human bone marrow cells with meayamycin B, a drug that affects mRNA splicing by inhibiting SF3B binding, similarly induced the formation of ringed sideroblasts in vitro. These data collectively suggest that the presence of ringed sideroblasts in individuals with RARS or RCMD-RS may be directly attributable to a reduction in SF3B1 activity within their erythroid precursor cells.

**Therapeutic Potential of Drugs That Target the Spliceosome**

To identify compounds with therapeutic potential—that selectively remove or normalize those cells that are positive for a spliceosome mutation—the effects that these mutations have on transcript processing will need to be more fully understood. Pharmacological intervention might prove particularly challenging if these mutations encode inactive proteins, or proteins that have dominant-negative activity. If they provide a gain-of-function, however, splicing patterns could potentially be normalized by exposure to one of the already identified compounds that inhibit spliceosome function.

In a screen of natural compounds to identify those with antitumorigenic properties, three bacterial fermentation products (FR901464, herboxidiene, and pladienolide) were identified as cytotoxic to multiple tumorigenic cell lines (42). Synthetic analogs with improved stability and solubility, such as meayamycin, E7107, and spliceostatin A (SSA), were subsequently developed. These drugs all inhibit spliceosome function by specifically targeting the SF3B protein complex, although it is unclear whether they target SF3B1, SF3B3, or the interface between these subunits (43,44). Studies suggest that E7107 and SSA both destabilize U2 snRNP assembly at 3’ splice sites by impairing the ability of SF3B to bind RNA (42); E7107 induces a conformational shift in the U2 snRNP that prevents its interaction with the branchpoint sequence, whereas SSA induces it to bind to “decoy” sequences that can occur upstream of the branchpoint sequence. However, it has been suggested that these drugs do not globally inhibit splicing at therapeutic concentrations, but instead perturb the expression of one or more proteins essential to the viability of a particular tumor cell type (42). An example of this phenomenon was recently observed in two non–small cell lung carcinoma cell lines that lack a spliceosome mutation (45). Meayamycin B treatment of these cells induced cell death by altering the relative amounts of the short and long isoforms of MCL-1, while not affecting the isoforms of another regulator of apoptosis, BCL-X. Although meayamycin B acted as an overall inhibitor of MCL-1 pre–mRNA processing, it predominantly altered the pattern of exons included in the mature transcript. This reduced the relative level of the anti-apoptotic MCL-1, isoform, with a concomitant increase in the level of the proapoptotic MCL-1, isoform, triggering cell death. This effect could be recapitulated by SF3B1 knockdown in these cell lines, confirming that meayamycin B treatment targeted the SF3B3 complex (45).

Although they result in tumor cell death in vitro, one concern is that the SF3B inhibitors may mimic or even exacerbate the effects of a disease-associated SF3B1 mutation. An example of this possibility was outlined earlier in this text: normal erythroid precursors differentiate into ringed sideroblasts following exposure to meayamycin B (41). The activity of these inhibitors will need to be tested on hematopoietic cells from patients with an SF3B1 mutation to clarify this issue. Another concern is likely to be drug resistance: an R1074H mutation in SF3B1 has already been identified and shown to confer pladienolide B resistance to colorectal carcinoma cells in vitro (46). Mutations at this residue have not been detected thus far in patients, but nevertheless have the potential to arise and be selected for during treatment, impairing the in vivo efficacy of this compound. As R1074 is located within the fifteenth HEAT repeat and the disease-associated SF3B1 mutations occur in other HEAT repeats (mainly in repeats 5–8; Figure 2), cells carrying an SF3B1 mutation might also have an impaired response to pladienolide B in vitro or in vivo.

An alternative strategy might be to employ natural regulators of splicing to correct the effect of any mutation.
The SR proteins provide an ideal target, as several forms of post-translational modification modulate their activity in healthy cells. This includes arginine methylation, which influences subcellular localization; acetylation of lysine residues, which may influence the stability of SR proteins; and phosphorylation of serine residues in the RS domain. Phosphorylation promotes spliceosome assembly by inhibiting any random binding to RNA and by facilitating interactions between SR proteins, and also influences alternate splicing patterns. SR proteins can be phosphorylated by topoisomerase I, and members of the SRPK (SR protein kinase) and CLK (Cdc2-like kinase) families. A number of small molecules are known to inhibit SR protein phosphorylation, including two that are already approved for the treatment of other conditions. Amiloride, an epithelial Na+ channel blocker used to treat hypertension, decreases SRSF1 phosphorylation (35), and chlorhexidine, a topical antiseptic, inhibits the activity of CLK family members (21). Other molecules that may have therapeutic potential include TG003, an inhibitor of CLK1 and CLK4 (27); SRPIN340, an SRPK1 and SRPK2 inhibitor (31); and NB-506, which inhibits the kinase activity of topoisomerase I (37).

A third potential approach involves altering the processing of key pre-mRNAs by targeting the regulatory sequences (ISEs and ESEs, or ISSs and ESSs) that modulate their differential splicing patterns. Modified antisense oligonucleotides could be used to modulate the function of these elements, increasing production of a beneficial protein isoform or decreasing a deleterious isoform. This approach has proved successful in a murine model of spinal muscular atrophy, a neurodegenerative disorder resulting from the inheritance of mutations that reduce the level of functional survival motor neuron (SMN) protein (47). Blocking oligonucleotides were generated against an ISS that prevents the inclusion of exon 7 in the mature SMN2 transcript; treatment of affected mice with these oligonucleotides corrected SMN2 splicing, restoring SMN expression in motor neuron cells and alleviating the disease phenotype. Phase I trials of these oligonucleotides in patients with spinal muscular atrophy are now under way. The synthesis of proteins with one or more isoforms that support tumor cell proliferation, such as MCL-1, BCL-X, PKM (Figure 3B), and VEGF-A, could be targeted in this fashion.

Are There Splicing Defects in Patients Lacking the Spliceosome-Associated Mutations?

Although a proportion of patients with MDS or CLL have acquired mutations in components of the pathway responsible for primary transcript processing, the majority of patients are unaffected. To date, studies that have not used an unbiased whole genome or exome sequencing strategy to detect spliceosome mutations, but have rather employed a targeted sequencing approach, have not looked at the entire repertoire of genes that encode spliceosome proteins. It can therefore be anticipated that a subset of patients who carry spliceosome mutations has not yet been identified. Do some or all of the remaining patients have any abnormalities in primary transcript processing? The detection of CPSF2, DDX3X, and XPO1 mutations in patients with CLL who lack a spliceosomal mutation (22) support the hypothesis that perturbation of mRNA processing may play an important role in the initiation or progression of this disease, as CPSF2 participates in mRNA polyadenylation, and DDX3X and XPO1 are both involved in the nuclear export of transcripts (48–50). CLL patients may also acquire mutations in EIF4A3 or MAGOH, two of the constituents of the exon junction complex (51). This complex forms at the junction between adjacent exons following spliceosome-mediated excision of the intervening sequence (Figure 1), and plays several roles in posttranscriptional processing, including nuclear export.

Constituents of other complexes involved in RNA processing, including the retention and splicing, cap-binding, and PRP19 complexes, may also be mutated (51).

Two recently published knockout mouse strains with a phenotype similar to human MDS suggest that patients with this disorder may also carry mutations that alter the correct processing of biologically important transcripts. One of these studies involved the deletion of Dicer, which encodes a microRNA (miRNA) endonuclease, from the osteoprogenitor cells present within the bone marrow microenvironment (52). This resulted in the emergence of an MDS-like phenotype, with mice developing anemia, leukopenia, and thrombocytopenia, and marked dysplasia in megakaryocytic cells. In contrast, loss of Dicer specifically from HSCs increased their number by altering the proportion undergoing apoptosis (53). The cell-extrinsic nature of the defect in Dicer-deficient osteoprogenitors was confirmed when the myelodysplastic phenotype was recapitulated by transplanting wild-type hematopoietic cells into Dicer-deficient recipients (52). Inactivation of DICER may not contribute to MDS development in humans, as DICER mutations have not been detected. However, there is an intriguing connection between DICER deficiency and Shwachman-Diamond-Bodian syndrome, a bone marrow failure syndrome that includes myelodysplasia among its features and results from the inheritance of a mutant SBDS allele (54). Gene set enrichment analysis of Dicer-deficient osteoprogenitor cells revealed that Shds transcription was dramatically reduced, and targeted deletion of Shds in these cells resulted in leukopenia and dysplastic changes in megakaryocytes and granulocytes (52). Acquired SBDS mutations do not occur in patients with de novo MDS, although reduced expression of SBDS, as well as DICER and DROSHA, another microRNA endonuclease, has been noted in mesenchymal cells obtained from these individuals (55). The level of each protein in mononuclear cells collected from these patients and from healthy donors was comparable, however, suggesting that unidentified factors in the marrow microenvironment may contribute to the overall disease phenotype. Studies are required to determine whether acquired mutations occur within the bone marrow stroma of MDS patients, and if so, to identify what their specific role is in the development of these disorders.

The second murine model linking abnormal splicing to the development of MDS is the Crebbp heterozygous mouse (56–58). It has long been appreciated that patients with therapy-related MDS may carry a translocation between chromosomes 11 and 16, which fuses together the MLL and CREBBP genes (59). The carboxy portion of CREBBP, which includes its bromodomain and histone acetyltransferase domain, is retained in the resulting chimeric protein, but the domain which binds nuclear hormone receptors is invariably lost. Mll-Crebbp expression in a murine transduction/transplantation model causes a myeloproliferative
phenotype that undergoes eventual leukemic transformation (60). In contrast, Crebbp-haploinsufficient mice initially develop a progressive pancytopenia (56). Reduced Crebbp levels do not influence HSC formation during embryogenesis nor perturb their ability to differentiate in vitro, but affect their ability to self-renew (57). Subsequent analyses revealed that the hematological disorder in these mice has myelodysplastic and myeloproliferative features, and is therefore an ideal animal model for studying human MDS/MPN overlap syndromes (61). Whereas the MDS-like phenotype is intrinsic to hematopoietic cells (V.I. Rebel, unpublished data), the myeloproliferation is a consequence of reduced Crebbp levels in the stromal and/or endothelial cells of the marrow microenvironment (58). Over time, about 40% of Crebbp-haploinsufficient mice spontaneously develop an acute leukemia that is most often myeloid in nature, with the loss of the remaining wild-type Crebbp allele accompanying disease transformation (56).

Microarray-based analysis of the transcriptome of HSCs from the fetal liver of Crebbp+/− mice and their wild-type littermates revealed that there are major differences in gene expression in these populations (62,63), as one might expect given the multiple roles CREBBP plays in coordinating and regulating gene transcription. Strikingly, more than half of sequences downregulated in Crebbp-haploinsufficient cells were intrinsic in origin, suggesting that they had smaller pools of unprocessed primary transcripts (62). Among the genes affected were those that encode proteins involved in RNA processing, including two SR proteins, Rbm39 and Sfsl2 (also referred to as Srekl1); Prp40A; Dgcr8, a double-stranded RNA binding protein that facilitates miRNA processing by Drosha and Dicer (64); and Musashi-2, an RNA binding protein whose overexpression alters HSC function in vivo and enhances the leukemogenicity of the BCR/ABL oncprotein (65). Another downregulated target was Malat1, a large noncoding RNA whose abundance is regulated by Dgcr8 and Drosha, and in turn regulates RNA splicing by modulating SR protein phosphorylation (64). These data collectively suggest that mutations that decrease CREBBP activity in hematopoietic cells substantially alter transcript processing within these cells, thereby contributing to the development of MDS.

Conclusions

With the advent of DNA sequencing technologies that for the first time allow rapid interrogation of the human genome, remarkable progress has been made in our understanding of the molecular pathogenesis of many malignancies. Of particular interest are the acquired mutations that affect components of the human spliceosome. Clinically relevant mutations that alter transcript splicing have been known for many years; estimates suggest that 15% of disease-causing mutations are located in the 5′ or 3′ splice sites, and that more than 20% of missense mutations lie in predicted splicing enhancer or silencer elements (http://www.dbass.org.uk). However, it was not appreciated until 2011 that disease-associated mutations might also target components of the splicing machinery. The identification of these mutations not only in patients with a hematological malignancy, but also in an expanding list of solid tumors, suggests that they play a fundamental role in establishing or maintaining the malignant phenotype. Drugs that normalize the changes in splicing associated with these mutations, or that target for degradation those cells bearing a mutation, should therefore be considered a priority.

References


Funding
This work was supported by funds from the University of Queensland Diamantina Institute (to LMS) and the Greehey Children's Cancer Research Institute (to VIR).

Notes
The authors of this manuscript take sole responsibility for the analysis and interpretation of the literature, writing of the manuscript, and the decision to submit the manuscript for publication. The authors have no conflicts of interest.

Affiliations of authors: Diamantina Institute, and Faculty of Health Sciences, School of Medicine, University of Queensland, Brisbane, Queensland, Australia (LMS); Translational Research Institute, Brisbane, Queensland, Australia (LMS); Greehey Children's Cancer Research Institute, Cancer Therapy and Research Center, and the Department of Cellular and Structural Biology, University of Texas Health Sciences Center at San Antonio (VIR).