Bioscience Reports (2023) **43** BSR20220149
https://doi.org/10.1042/BSR20220149

PORTLAND PRESS

Review Article

How does precursor RNA structure influence RNA processing and gene expression?

Austin Herbert, Abigail Hatfield and Lela Lackey

Department of Genetics and Biochemistry, Clemson University Center for Human Genetics, Greenwood, SC, U.S.A.

Correspondence: Lela Lackey (lelal@clemson.edu)

RNA is a fundamental biomolecule that has many purposes within cells. Due to its single-stranded and flexible nature, RNA naturally folds into complex and dynamic structures. Recent technological and computational advances have produced an explosion of RNA structural data. Many RNA structures have regulatory and functional properties. Studying the structure of nascent RNAs is particularly challenging due to their low abundance and long length, but their structures are important because they can influence RNA processing. Precursor RNA processing is a nexus of pathways that determines mature isoform composition and that controls gene expression. In this review, we examine what is known about human nascent RNA structure and the influence of RNA structure on processing of precursor RNAs. These known structures provide examples of how other nascent RNAs may be structured and show how novel RNA structures may influence RNA processing including splicing and polyadenylation. RNA structures can be targeted therapeutically to treat disease.

Introduction

Are precursor RNAs structured?

The sequence of an RNA influences its biological activity. Sequence information still predominates as the most studied aspect of a nucleic acid. However, due to its single-stranded and flexible nature, RNA naturally forms structures as soon as it is synthesized by an RNA polymerase, reviewed in [1,2]. There is clear evidence for robust and reproducible RNA structures in RNA transcripts, both in human cells and several model organisms [3–8]. RNA structures can range from stable, consistent folds to flexible structural ensembles, reviewed in [9]. Transcriptome-wide structure analysis has revealed patterns of functional structure in processed RNAs, including flexible structural transitions around start and stop codons [3,5,6,8]. However, little is known about RNA structure in nascent human RNAs before they undergo processing to become mature transcripts. There is evidence that RNA structure is altered at different stages of an RNA molecule's life cycle. Liu et al. found major differences between nuclear and cytoplasmic RNA structures in *Arabidopsis*, which suggested that RNA structures changed significantly from the nascent RNA to the mature transcript [10]. Structures in 3′UTRs can vary during different stages of development in zebrafish [11]. In yeast, there are differences in structure between the same RNAs *in vivo* versus RNAs extracted from the cell [4]. Direct study of nascent RNA structure is important to understand the relationship between precursor and mature RNA structures and how precursor RNA structure can impact RNA processing and gene expression. Although still limited, several studies have identified transcriptome-wide patterns of RNA structures in nascent RNAs [10,12].

What methods are available to study nascent RNA structures?

Experimentally based secondary structural models of RNAs can be created with enzyme and chemical probing data, reviewed in [13]. In particular, chemical probing combined with next-generation sequencing has become very popular due to its ability to generate data on long RNAs and multiple RNAs at the same

Received: 20 June 2022
Revised: 17 January 2023
Accepted: 23 January 2023

Accepted Manuscript online: 23 January 2023
Version of Record published: 01 March 2023

Bioscience Reports (2023) **43** BSR20220149
https://doi.org/10.1042/BSR20220149

Table 1 Chemical probing methods

Protocol	Brief description	Examples	Pros	Cons
Reverse transcriptase (RT)-stop	During reverse transcription chemical adducts cause RT fall off. Truncated products are run on a gel.	DMS RT-Stop [179,180]	Capable of measuring nucleotide accessibility in flexible RNAs	Restricted to short sequences, one RNA at a time, does not determine specific base-pairs
MaP (mutational profiling)	During reverse transcription chemical adducts are replaced with mutations and read out by sequencing.	DMS-MaP [14], SHAPE-Map [15,181]	Capable of measuring accessibility in long, flexible RNAs and multiple RNAs at the same time	Requires high read depth, does not determine specific base pairs
RNA pulldown	The chemical probe is tagged (i.e., click chemistry, biotin-conjugation, etc.). RNAs with adducts are enriched.	icSeq [8], SHAPES [182]	Captures low abundance RNAs by enrichment	Enrichment may disrupt reactivity calculations, does not determine specific base pairs
Protein immunoprecipitation	Probed RNAs are co-precipitated by a protein-targeted antibody for analysis.	tNET-Structure-Seq [12], fSHAPE [183,184]	Specifically targets RNAs associated with a protein	Requires high read depth, depends on antibody specificity and affinity, does not determine specific base pairs
Hybridization-capture	Probed RNAs are targeted by tagged oligonucleotides (i.e., biotinylated-U) for analysis.	SHAPE-MaP enrichment [31]	Can measure accessibility in low abundance RNAs	Requires production of oligonucleotides to target RNAs, does not determine specific base pairs
Cross-linking	Nucleotides in close proximity are covalently linked by UV and/or crosslinking compounds. The RNA is enriched, cleaved, and ligated for junction analysis.	PARIS [185], CLASH [186], SPLASH [187], LigR-Seq [188], RIC-Seq [189]	Identifies long-distance base pairing, determines specific base pairs	Can have false positives, requires high read depth
Selection by 3′ end sequencing	Probed RNAs undergo 3′ sequencing using polyA oligo hybridization, cleavage and polyA priming.	DIM-2P-Seq [60]	Improves structure definition at the 3′ end of transcripts	Requires high read depth, does not determine specific base pairs

Listed techniques are based on chemical probing where RNA is treated with a chemical that typically forms adducts with accessible nucleotides. The modified RNA is converted to DNA by reverse transcription.

time [14,15]. Chemical probes include those that react with the ribose backbone (selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) reagents) and those that react with nucleobases, such as dimethyl sulfate and carbodiimides [16,17]. There are a wide variety of chemical probing applications (Table 1). While chemical probing primarily reveals secondary structure, there are other structure modeling techniques capable of measuring tertiary structures including cryo-electron microscopy (cryo-EM), nuclear magnetic resonance (NMR) and small angle X-ray scattering (SAXS). Recent advances in cryo-EM technology have resulted in structures of large RNA and protein (RNP) complexes like the spliceosome and ribosome [18,19], suggesting that deriving tertiary structural models for stable structures in large pre-processed RNAs may be possible. Secondary and tertiary RNA structural models can be modelled computationally and incorporate experimental chemical probing data [20]. These calculated structures can be based on combinations of thermodynamic parameters, machine learning and experiments, including highly cited methods like RNAstructure, mFOLD and others [21–27]. However, though less time consuming than experimentally generated structures, the accuracy of *de novo* models of RNA structure, both secondary and tertiary, is questionable [28–30].

A major hurdle in acquiring experimental secondary structure data from *in vivo* systems is that most methods depend on a significant number of molecules to quantify reactivity (Table 1, Cons). Thus, current *in vivo* methods are limited to high abundance transcripts. Low abundance transcripts that do not meet copy number thresholds require *in vitro* or more elaborate techniques to deduce structure [31,32]. The low abundance of precursor RNAs is one reason that their structures are understudied. Some techniques are beginning to address the problem of low abundance RNAs, including enrichment for low abundance targets during chemical probing (Table 1, RNA pull-down and Hybridization capture) [31,33,34]. Another hurdle in studying the structure of precursor RNAs is the flexible nature of RNA molecules. Although some RNAs have stable structures, such as transfer RNAs and RNAs found in the spliceosome and ribosome, most RNAs have many possible structures that are energetically similar. This ensemble effect of RNAs must be considered when determining what structures are biologically relevant. The long length of most introns compounds this problem for precursor RNAs. Computational modeling to analyze ensembles is being applied to assist in this problem [35–37]. Additionally, many RNAs are likely to have long-distance and tertiary interactions

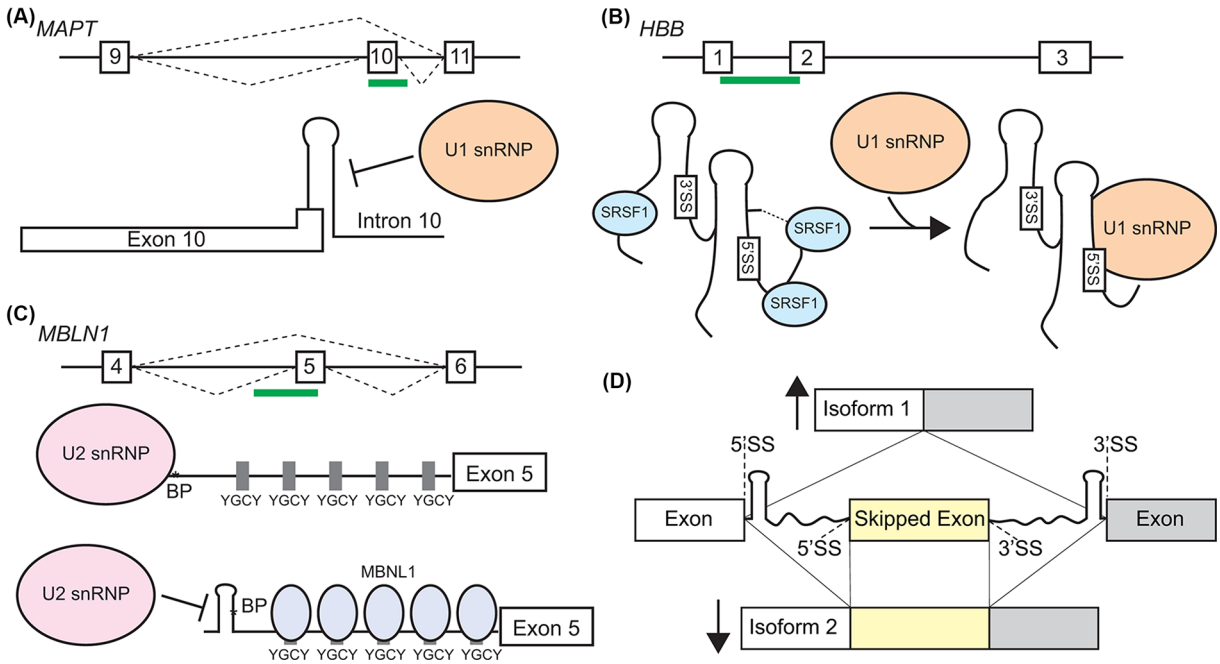

Figure 1. RNA structures influence precursor RNA processing

(**A**) Hairpin elements can block 5′ splice site recognition by interfering with U1 snRNP binding. *MAPT1* RNA exon 10 alternative splicing is controlled by hairpin structure at the 5′ splice site. (**B**) RNA structure can bring distal elements in close proximity. The global fold of the *HBB* RNA is mediated by SRSF1 binding and orients the 5′ and 3′ splice sites for U1 snRNP interaction and efficient splicing. (**C**) Recognition of RNA elements are control processing. MBNL1 protein binds to its own RNA. MBNL1 binding causes remodeling of the RNA structure around the branchpoint and represses exon 5 inclusion. (**D**) Transcriptome-wide structural analysis of nascent RNA found clear structural 'steps' in proximity to efficiently spliced exons (top). The structure 'steps' around frequently skipped exon are less evident (bottom). Cartoon depiction based on [12].

that may be functional but remain difficult to map [38,39]. Little is known about the importance of long-distance interactions in pre-processed RNAs. Experimental mapping techniques are available to document long-distance interactions (Table 1, Cross-linking). Finally, the nature of co-transcriptional processing adds a temporal element to precursor RNA structure modeling. Introns can be spliced out of order, making it difficult to predict which nucleotides are available for structural interactions during transcription of an RNA [40–42]. There are experimental and computational approaches that partially address the temporal nature of RNA folding by targeting temporally associated proteins (Table 1, Protein immunoprecipitation) [12,43]. However, the specialized approaches that address the problems of RNA structure modeling in dynamic, long and low abundance RNAs are not easily broadly applied.

Are structures in precursor RNAs functional?

There is no self-evident reason to assume that the structures that RNA molecules form are inherently functional. However, careful research has identified many functional RNA structures and their mechanisms. RNA structures can act to block recognition motifs. Short hairpin motifs that block recognition of the 5′ splice site are common and one of the earliest known structures to functionally impact splicing [44,45], reviewed in [46]. For example, the nascent *MAPT* (*microtubule associated protein tau)* RNA has a hairpin element at the 5′ splice site of exon 10 (Figure 1A). *MAPT* is important in neural biology and its precursor RNA is alternatively spliced to create at least 6 isoforms, reviewed in [47]. A splice junction in *MAPT* (exon 10–intron 10) is normally spliced at an equal ratio, resulting in mix of mature isoforms with either 3 or 4 microtubule binding domain repeats, which code for Tau proteins with different biological activity [48]. A hairpin at the *MAPT* exon10–intron10 junction directly overlaps with the 5′ splice site and can be disrupted by disease-associated variants [49,50]. The *MAPT* hairpin blocks normal spliceosomal recognition by the U1 snRNP at the 5′ splice site and causes exon skipping and formation of the shorter 3R *MAPT* mature transcript [51] (Figure 1A). RNA structure in pre-processed transcripts has been shown to block U1 interactions and alter splicing in other precursor RNAs in addition to *MAPT*, including *SMN2, VWF, ATM* and *BCL2L1* [52–56] (Table 2).

Bioscience Reports (2023) **43** BSR20220149
https://doi.org/10.1042/BSR20220149

Table 2 Mammalian genes containing functional RNA structures that affect RNA processing

Gene symbol	Gene name	Impact	Mechanism	Citations
MAPT	Microtubule associated protein tau	Splicing	Potential to block U1 snRNP binding	[50,51]
XBP1	X-box binding protein 1	Splicing	Recognition by IRE1	[12,190]
H2AC11, Histones	H2A clustered histone 11	3′ end cleavage	Recognition by SLBP	[129]
PLEC	Plectin	Splicing	Recognition by SNRPA1	[61]
TNNT2	Troponin T2, cardiac type	Splicing	MBNL1 and U2AF65 competition	[63,191]
MBNL1	Muscleblind-like splicing regulator 1	Splicing	Recognition by MBNL1	[31,62]
FN1	Fibronectin 1	Splicing	Recognition by SR proteins	[56]
SMN2	Surviver of motor neuron 2, centromeric	Splicing, polyadenylation	Block U1 snRNP binding, PIE element	[55,90,119]
VWF	von Willebrand factor	Splicing	Potential to block U1 snRNP binding	[53]
ATM	ATM serine/threonine kinase	Splicing	Potential to block U1 snRNP binding	[192]
CFTR	CF transmembrane conductance regulator	Splicing	Potential to interfere with U6 snRNP	[192]
TERT	Telomerase reverse transcriptase	Splicing	G-quadruplex RNA	[193,194]
TP53	Tumor protein p53	Splicing	G-quadruplex RNA	[195]
BCL2L1	BCL2 like 1	Splicing	G-quadruplex RNA and U1 blocking hairpin	[52,55]
FMR1	Fragile X messenger ribonucleoprotein 1	Splicing	FMR protein G-quadruplex binding	[196]
CD44	CD44 molecule (Indian blood group)	Splicing	G-quadruplex RNA	[197]
ENAH	ENAH actin regulator	Splicing	Long-distance pairing blocks RBFOX binding	[198,38]
DST	Dystonin	Splicing	Long-distance pairing	[38]
PLP1	Proteolipid protein 1	Splicing	Long-distance pairing	[199,38]
SF1	Splicing factor 1	Splicing	Long-distance pairing	[38]
DNM1	Dynamin 1	Splicing	Long-distance pairing	[38]
ATE1	Arginyltransferase 1	Splicing	Long-distance pairing	[38,200]
PSEN2	Presenilin 2	Splicing	Unknown	[201]
FGB	Fibrinogen β chain	Splicing	Promotes TRA2B binding and splicing	[202]
CENPB	Centromere protein B	Polyadenylation	Collapsed distance between polyA site and cleavage site	[60]
U1A	U1 small nuclear ribonucleoprotein A	Polyadenylation	PIE element	[118,155]

RNA structure can also function to collapse the distance within long sequences to bring RNA elements into close proximity. Some nascent RNAs may use structure to condense long intronic regions to form structures that are suitable scaffolds for early spliceosome recognition and activation. The human *AdML (adenovirus 2 major late transcript IVS1*) precursor RNA has a global fold important for splicing [57]. Disrupting the structure of pre-processed *AdML* RNA prevents it from being recognized efficiently by the U1 snRNP in *in vitro* studies. FRET analysis confirms that the 5′ and 3′ splice sites are in close proximity in the normal structure, but not in poorly spliced mutants with altered structures [57]. In the *AdML* precursor RNA the global fold is not dependent on proteins. Similar studies of the human *HBB (β globin*) precursor RNA support the role of global RNA structure in recruiting U1 and promoting splicing [58] (Figure 1B). However, the global fold of the pre-processed *HBB* RNA is influenced by binding of SRSF1 protein as a structural stabilizing factor (Figure 1B). The ability of RNA structure to influence splicing by collapsing the distance between the branchpoint and the 3′ splice site has been documented in several introns in yeast [59]. RNA structure may also collapse the distance between the polyadenylation recognition motif and the cleavage site during 3′ processing and polyadenylation of nascent RNAs [60].

RNA structures can be recognized specifically, often in combination with sequence elements. Small nuclear ribonucleoprotein polypeptide A′ (SNRPA1) recognizes intronic sequence-independent stem structures combined with

sequence-dependent loops to promote splicing of cassette exons in multiple genes, including *plectin* (*PLEC*) precursor RNA [61]. SNRPA1 splicing of *PLEC* contributes to a prometastatic cellular environment and is associated with progression and poor prognosis in breast cancer [61]. *MBLN1* precursor RNA is autoregulated by MBLN1 protein binding at primarily unpaired YGCY motifs close to the 3′ splice site [62] (Figure 1C). Binding of MBLN1 to *MBLN1* RNA restructures a distal branchpoint and results in exon 5 skipping and an isoform of MBLN1 protein with different subcellular localization [31] (Figure 1C). RNA structural elements can utilize multiple attributes of folding. In addition to containing structures that specifically bind MBLN1, the global fold of the *MBNL1* exon 5 has also been shown to bring the 3′ and 5′ splice sites into close proximity [31,62]. MBLN1 protein also binds to loop elements with YGCY sequences in other nascent RNAs, including cardiac troponin RNA (Table 2) [63].

Despite technical difficulties in performing transcriptome-wide studies of precursor RNA structure, recent studies have broadly analyzed precursor RNA structure to identify global patterns of functional base pairing. Saldi et al. found that there are higher-order structural 'steps' that demarcate efficiently spliced and structured introns from less efficiently spliced exons in human cells [12]. The researchers used tNET-Structure-Seq to acquire structural data on nascent RNAs in human cell lines. tNET-Structure-Seq combines enzymatic and chemical structure probing with RNA Polymerase II immunoprecipitation to determine the accessibility of nascent RNA nucleotides. Introns spliced co-transcriptionally were associated with clearer structural 'steps' at the 5′ and 3′ splice sites when compared with splice junctions in introns that were spliced post-transcriptionally [12]. Structural 'steps' are characterized by a disparity in nucleotide accessibility around the exon–intron junction. Typical structural 'steps' have higher accessibility on the exonic side of the junction and lower accessibility on the intronic side of the junction (Figure 1D). The magnitude of a structural 'step' also leads to splicing preferences in cassette exons with bigger differences between the accessibility of the exon and intron leading to more efficient splicing. For example, mutually exclusive exons, a type of cassette exon, are primarily spliced post-transcriptionally and exon-inclusion or exclusion is influenced by RNA structure. Excluded exons had more robust 'steps' at the farthest 3′ splice site, whereas included exons were more likely to have weak 'steps' at the farthest 3′ splice site [12]. These finding are consistent with the general lack of base-pairing at the 3′ splice site and structural differences based on splicing efficacy found in *Arabidopsis* seedlings [10,64]. The pattern of unpaired, accessible exonic nucleotides and paired intronic nucleotides in efficiently spliced transcripts has also been found in mouse precursor RNAs [65] and other organisms [31,57,64–67], reviewed by [68,69].

Approaches to identifying functional structures

How can we distinguish functional structures in a sea of RNA structure?

All RNAs have structural patterns, many of which are functionally important, yet it can be difficult to identify which RNA structures are relevant without prior knowledge. Unlike many nascent RNAs, the survival of motor neuron 2 (*SMN2)* precursor RNA has been extensively mapped for structural elements [55,70–72], reviewed in [73]. *SMN2* splicing studies have focused on exon 7, which can either be skipped or included. Two terminal stem-loops (TSL1 and TSL2) within exon 7 and several structures within intron 7 influence splicing (Figure 2A) [55,70,71,74]. These structures block U1 recognition of the exon 7 at the 5′ splice site and influence RNA-binding protein (RBP) interactions with enhancer and silencer functions. Inclusion of exon 7 allows *SMN2* mature transcripts to produce active SMN protein and promote neural survival in the event of defects at the *SMN1* locus [75,76]. Similar to exon 7 in *SMN2*, exon 3 of *MCL1* can be skipped or included, resulting in the production of short (*MCL1-S)* and long (*MCL1-L)* RNA and protein isoforms with either pro- or anti-apoptotic functions [77]. In contrast with the extensive research into *SMN2* RNA splicing, very little is known about how RNA structure influences RNA processing in most other nascent RNAs, including *MCL1*. We use *SMN2* and *MCL1* nascent RNAs as examples to discuss how to identify functional nascent RNA structures for validation.

One approach to identifying functional regions that may have important RNA structures is to take advantage of conservation data [78,79]. Highly conserved regions within a gene may have functional importance. Both *SMN2* (Figure 2B,C) and *MCL1* (Figure 2F,G) follow a typical conservation pattern with higher conservation values in the exonic regions and lower conservation in the intronic sequences. A region of SMN2 that stands out is the continuation of high conservation scores into the intronic region of intron 7 (Figure 2B,C, boxed). This region overlaps with TSL2 and internal stem loop 1 (ISL1), which sequester the 5′ splice site and promote exon 7 skipping [55]. Additional conserved regions overlap with the functional long-distance internal stem 2 (ISTL2) [71] and TSL4. In *MCL1*, high conservation values are also extended into intron 2, overlapping with the 3′ splice site and experimentally identified branchpoints [80] (Figure 2F,G, boxed). In addition to standard conservation data, covariation analysis can be used to identify functional structures by looking for regions where base pairing is conserved [81,82]. Covariation has been important for identifying functional structures in lncRNAs; however, it is not commonly detected in human protein

Bioscience Reports (2023) **43** BSR20220149
https://doi.org/10.1042/BSR20220149

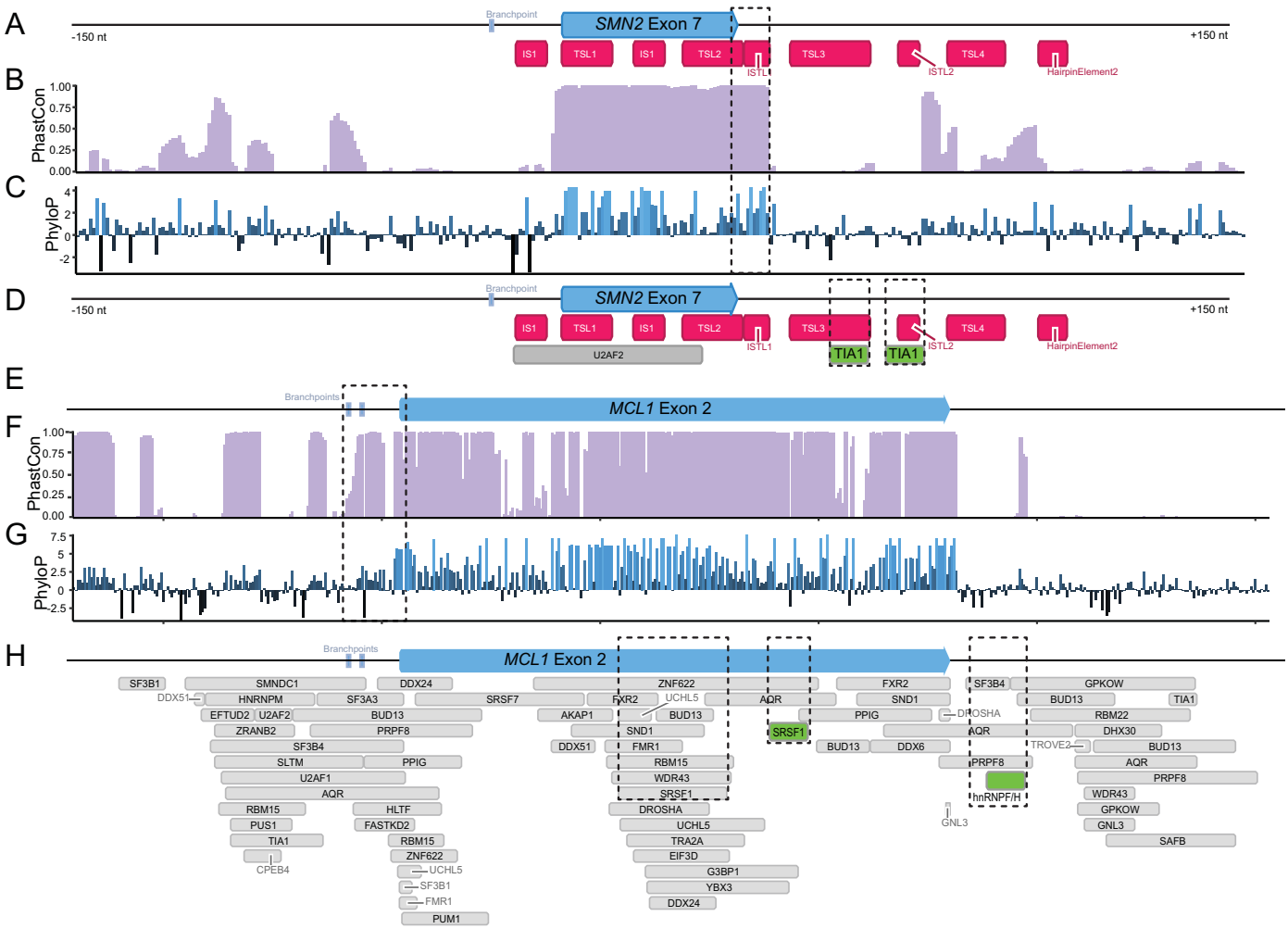

Figure 2. Identifying potential functional structures in *MCL1* RNA using nucleotide conservation and protein binding sites

(**A**) Schematic of human *SMN2* exon 7 surrounded by 150 intronic nucleotides on both sides. Known RNA structures in *SMN2* are annotated (red). (**B** and **C**) Nucleotide conservation in *SMN2*. Higher values indicate more conservation. A region of high conservation extends into the 5′ splice site of exon 7 and overlaps with TSL2 and ISTL1 (boxed). (**D**) Protein binding sites across *SMN2* from ENCODE (gray) and published studies (green) are mapped on to the schematic. Binding sites for the RBP TIA1 overlap with TSL3 and ISTL2 structures (boxed). (**E**) Schematic of human *MCL1* exon 2 surrounded by 150 intronic nucleotides on both sides. (**F,G**) Nucleotide conservation in *SMN2*. Higher values indicate more conservation. A region of high conservation extends into the 3′ splice site of exon 2 and overlaps with the branchpoint region (boxed). (**H**) Protein binding sites across *MCL1* from ENCODE (gray) and published studies (green). Binding sites for regulatory RBPs, SRSF1 and hnRNPF/H are indicated (boxes). For both RNAs, schematics were visualized with Geneious Prime v2022.1.1 [203]. Branchpoints were annotated based on experimental data [80]. Conservation data were retrieved from UCSC table browser (Cons 100 Verts, phastCon, phyloP100way) [79]. ENCODE eCLIP data were retrieved as bigBed narrowPeak annotations filtered by a *P*-value of < 0.05, and annotations collapsed for overlapping peaks of the same protein [85]. All data reference hg38.

coding RNAs [83,84]. Conservation data is not structure-specific, however, in *SMN2* conservation data indicates important structural regions, suggesting that highly conserved regions in *MCL1* may have structural significance.

Many nascent RNA processing steps rely on RBP interactions. Mapping RBP-binding sites onto RNA can highlight important regulatory regions. Structures nearby or that overlap with RBP sites may influence protein interactions. Experimentally determined RBP binding sites are available in published enhanced cross-linking and immunoprecipitation (eCLIP) databases, such as the ENCORE dataset in ENCODE [85–87]. In the ENCORE dataset, *MCL1* has many RBPs bound to intronic and exonic regions, including SRSF1, which is known to affect *MCL1-S* to *MCL1-L* isoform ratios [88] (Figure 2H, boxes). The ENCODE SRSF1 binding site overlaps with a hotspot of RBP binding where 15 additional RBPs have been identified at the same position. The structures of hotspot elements may influence competitive RBP binding. Studies of *MCL1* suggest that hnRNPF/H regulates *MCL1* splicing and binds within intron 2 [89] (Figure 2H, boxed). However, each individual study is limited to select tissues and time-points. *SMN2* is a neural factor and is not expressed in the ENCODE project cell lines. Only one RBP, U2AF2 is significantly associated with *SMN2* in this dataset. However, experimental studies on *SMN2* have been performed in multiple laboratories suggesting that more than 40 RBPs bind pre-processed *SMN2* RNA around exon 7 and influence its splicing, reviewed in [90]. For example, the RBP TIA1 promotes *SMN2* exon 7 inclusion by binding with intron 7 and recruiting the U1 complex in proximity to the 5′ splice site [91]. TIA1 binding overlaps with the intronic TSL3 structure (Figure 2D, boxed). Antisense oligonucleotides targeting *SMN2* near the TIA1 site are predicted to open TSL3, make the TIA1 binding site more accessible, and promote exon 7 inclusion [72]. Predicting RBP sites is another option for genes that are less studied than *SMN2* but not expressed in commonly used cell lines [92]. As we learn more about the preference of different RBPs for sequence and structural features we will be able to better predict the impact of RNA structures within RBP-binding sites. Since protein binding sites can be influenced by RNA structure, they are important regions for structural studies.

Another way to identify functional regulatory elements that may affect RNA processing is to map disease-associated variants onto the RNA, such as variants found through GWAS and familial studies and in databases such as ClinVar and HGMD [93,94]. Additional variant-based approaches can use quantitative trait loci (QTLs), which are variants associated with a variety of phenotypes like expression (eQTLs), splicing (sQTLs) and alternative polyadenylation (apaQTLs) [95–97]. However, there are generally few disease-associated variants mapped to genes, particularly in non-coding regions such as introns. In *MCL1* and *SMN2* there are only a handful of disease-associated variants (Figure 3B,E). Rare variants and somatic mutations may also be informative for identifying functional regions of an RNA [98,99]. To directly connect variants with RNA structure elements there are computational models to estimate the impact of a variant on local and global RNA structure [100–103], reviewed in [104]. Variants that change RNA structure are relatively common and are called riboSNitches [3,100,105], reviewed in [106]. We find that across *SMN2* there are 95 riboSNitches, while in *MCL1* there are 40 riboSNitches [102] (Figure 3B,E). RiboSNitches that overlap with rare or phenotypic variants are candidates that indicate functional structures that may be involved in phenotypic differences. In *SMN2* a disease-associated variant found in ClinVar is also predicted to be a riboSNitch (Figure 3B, arrow). This riboSNitch falls within IS1, which is known to have structure that regulates *SMN2* splicing [70]. Another riboSNitch in *SMN2* overlaps with a somatic variant from COSMIC in the regulatory *SMN2* intron 7 hairpin element 2 (Figure 3B, arrow) [74]. Similar somatic variants are predicted to be riboSNitches in *MCL1*, including a riboSNitch in close proximity to a ClinVar variant (Figure 3E, arrows).

Finally, rather than starting from functional elements and analyzing whether RNA structure may impact that function, we can also start from experimental or predicted structural models. There are many webservers and software packages that can predict minimum free energy models or base-pairing probabilities with reasonably high expected accuracy [21–26]. Within these structural models, highly probable hairpin/stem-loop motifs are common regulatory elements, reviewed in [107]. In *SMN2*, selection of strong hairpin elements from an experimentally based structural model for further study would have identified internal stem 1 (IS1), TSL2 and TSL3, all of which have been shown to structurally influence *SMN2* splicing (Figure 3C) [55,70]. These elements are predicted despite the limited region selected for folding. In addition to these known structures in *SMN2* there are two highly probably hairpin motifs within *SMN2* intron 6 that could influence processing of *SMN2* precursor RNA (Figure 3C). Structural models of *SMN2* have been used by others to predict novel functional elements in *SMN2* [73]. *MCL1* precursor RNA also has highly probable hairpin and nested hairpin motifs in an experimentally based structural model (Figure 3F). In the nested hairpin, base-pairing is predicted between the branchpoint region and exonic sequence suggesting that this structure could affect 3′ splice site recognition and contribute to regulatory processing of *MCL1* RNA.

Currently it is difficult to identify functional structures within a precursor RNA. *SMN2* RNA is a well-studied model that shows the ability of conservation data, RBP binding site analysis, variant mapping and structural models to identify RNA structures that many influence *SMN2* splicing. We also highlight regions within the understudied

Bioscience Reports (2023) **43** BSR20220149
https://doi.org/10.1042/BSR20220149

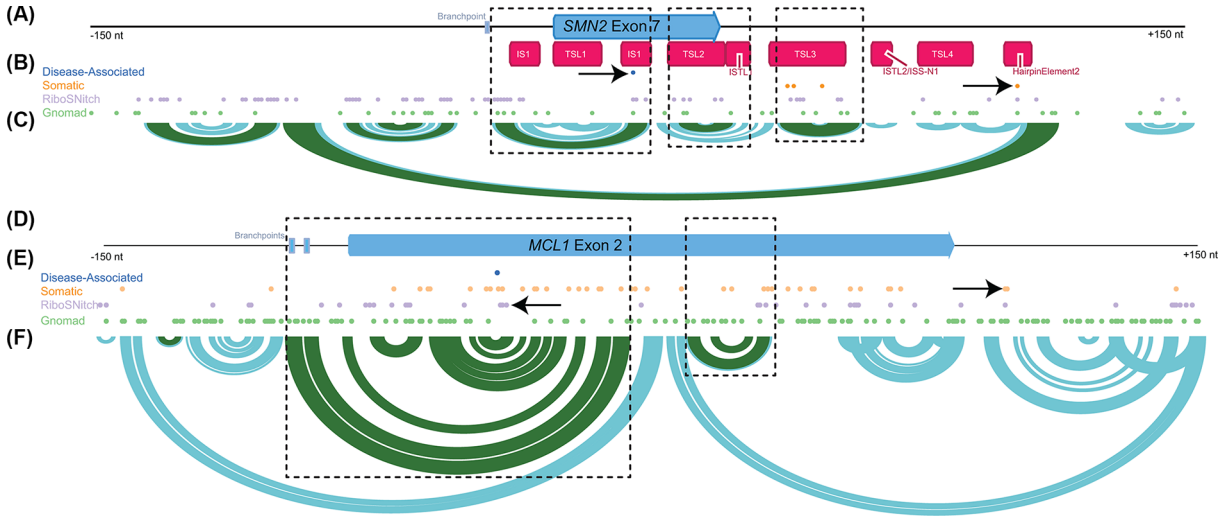

Figure 3. Identifying potential functional structures in *MCL1* RNA using genomic variation and RNA structure models

(**A**) Schematic of human *SMN2* exon 7 surrounded by 150 intronic nucleotides on both sides. Known RNA structures in *SMN2* are annotated (red). (**B**) Genomic variants in *SMN2* from a variety of sources including variants that are disease-associated (blue), somatic (orange), predicted to change RNA structure (riboSNitches, purple) and inherited (green). The ClinVar disease-associated variant in SMN2 is a riboSNitch (arrow). Likewise, a somatic mutation in Hairpin Element 2 is a riboSNitch (arrow). (**C**) Arc diagram of the *SMN2* RNA structure generated with published SHAPE-MaP data [204] showing highly probable base pairing (>80%, green) and moderately probable base pairing (30–80%, blue). Predicted structures overlap with published structural elements (boxes). (**D**) Schematic of human *MCL1* exon 2 surrounded by 150 intronic nucleotides on both sides. (**E**) Genomic variation in the *MCL1* exon 2 region, colored as indicated above. RiboSNitches in proximity to the ClinVar disease-associated variant overlap with somatic mutations (arrow). We also highlight a riboSNitch that is a somatic mutation (arrow). (**F**) *MCL1* precursor RNA structure base-pairing probabilities generated from *in vitro* 5NIA SHAPE-MaP data and analyzed with shapemapper2 [205]. For both RNAs variants were retrieved from gnomAD [99], COSMIC [98], RNAsnp screening [102], and ClinVar [93]. RNA structures were generated with the RNAStructure package functions (partition and ProbabilityPlot [206]) and visualized in IGV [207].

MCL1 RNA that may be of interest structurally to understand *MCL1* splicing. As only a handful of human genes have clear precursor RNA structural annotation, the lack of known functional structures even in a highly regulated gene like *MCL1* is unsurprising. The field will continue to acquire better secondary and tertiary structure models for *SMN2* and *MCL1* as new technologies for RNA structure modeling emerge, such as chemical probing techniques (Table 1) and cryo-EM. Although individual gene annotation for functional RNA structures is slow, an important goal is to identify enough functional structures in nascent RNAs to accurately predict functional structures directly from sequence.

Mechanisms underlying structural recognition and dynamics

How do precursor RNA structures influence RNA-binding protein interactions?

Proteins are principal effectors of cellular function. Proteins that interact with RNA are common and many studies have explored the sequence and structural preferences of RBPs [108–111]. The interaction between RBPs and RNA can be understood as a modular interaction between RNA-binding domains (RBDs) in the protein and target RNA elements. There are 16 well known RBDs, and likely additional non-canonical domains, reviewed in [112]. RBDs can be repeated or varied within a protein and the spacing between RBDs can influence how the protein interacts with its target RNA. Interactions between RBDs and RNA elements are governed by basic physical principles: hydrogen bonding, electrostatics, and base-stacking, reviewed in [113]. All three of these mechanisms are used by most RBDs for structure and sequence specific interactions. Most RBPs have degenerate sequence preferences that favor structurally accessible RNA sequences [108]. In general, sequence-specific interactions between RBPs and RNA occur through hydrogen bonding. Hydrogen bonds cannot readily form when the target nucleotides are base paired. However, even

though most RBPs target unpaired nucleotides for sequence specificity, structural context does influence RBP interaction. This allows RBPs to differentiate between RNA binding elements even when sequence motifs are degenerate or similar to other RBP binding sites [108].

One of the most common types of RBD is the RNA recognition motif (RRM). The U1A/SNF/U2A″ family is an example of RRM-containing proteins that can recognize structured RNA [114]. U1A is part of the spliceosome U1 snRNP. It binds specifically to the U1 stem-loop II (SLII) through structure specific base-stacking (Figure 4A, maroon), electrostatic interactions (Figure 4B, pink) and sequence-specific hydrogen bonding interactions (Figure 4A, orange) [115,116]. U1A binding is very specific and discriminates between stem-loops in U1 and U2 RNAs [117]. Despite this specificity, U1A is also capable of binding PIE RNA elements in its own 3′UTR and several other RNAs, including *SMN2*, to regulate polyadenylation [118,119] (Table 2). PIE elements are similar to the U1 SLII at the structure and sequence levels. Within a PIE element, duplicated stem-loops dimerize U1A (Figure 4C) and directly interact with polyA polymerase [120–122]. Although U2B contains a leucine-rich RBD rather than an RRM like U1A, it has a similar multifunctional ability. U2B interacts with the U2 RNA stem-loop IV with both structure and sequence specificity. However, in cancer cells, U2B has an extra-spliceosomal function wherein U2B binds intronic stem-loop structures and promotes splicing of cassette exons [61].

Double-stranded RNA-binding domains (dsRBDs) generally have strong structural preferences for their target RNAs. The cleavage factors Drosha and Dicer both contain dsRBDs and interact with precursor microRNAs through structure-dependent mechanisms [123,124]. MicroRNAs are processed from precursor RNAs (pri-miRNAs) that form extended stem-loops and are cleaved into pre-microRNAs and finally mature miRNAs. Pri-miRNAs interact with the microprocessor complex containing Drosha based on their helical RNA structure [125,126]. Drosha-dsRNA interactions are mediated by electrostatic interactions between the dsRBD (Figure 4D, gray) and two minor grooves, with minimal major groove interactions [127,128]. The extended stem-loop structure of pri-miRNAs is recognized at the transition between unpaired and paired segments at the bottom and top of the stem region. Flexibility in the pri-miRNA stem alters Drosha cleavage, ultimately resulting in the production of isoforms of mature miRNAs that can target different transcripts for translational repression [123]. The RBP Drosha and its partner DGCR8 recognize these structural junctions and structural changes in pri-miRNA, partially mediated by DDX3X, that result in different Drosha processing [123]. In addition to structure-specific recognition, Drosha also recognizes specific nucleotide sequences that are important for processing of pri-miRNAs. After Drosha processing, Dicer cleaves pre-miRNAs into duplex miRNAs. Dicer also recognizes structural elements of the pre-miRNA to discriminate between pre-miRNAs and other classes of RNA to cleave true pre-miRNAs into mature miRNAs [124].

While many RBPs have been extensively studied and their specificity determined, there are more than 2000 predicted RBPs in the human genome, and many 'moon-lighting' proteins with unrecognized RNA binding potential, reviewed in [112,113]. High-throughput studies to broadly characterize RNA-binding protein characteristics demonstrate the importance of both sequence and structural recognition [108,109]. Although the majority of RBPs prefer unpaired sequence motifs, they display varying sensitivity to structure. For example, RBPs including RBM22, RBM6, PRR3 and BOLL display preferences for motif-based partial pairing or a motif surrounded by paired nucleotides [108]. The little-known RBP ZNF326 even prefers sequence-based recognition within a completely paired structural motif [108]. By varying their ability to recognize structured sequences, RBPs effectively create a binding preference by incorporating the surrounding structural context of short sequence motifs [108]. Additional studies on the binding preferences of RBPs can help determine how RBPs recognize RNA and regulate gene expression.

How is precursor RNA structure modulated in the cell?

While the sequence of an RNA is the primary input for many computational structure models, biological regulation affects the structure of RNA in the cell. One mechanism of regulation is the speed of transcription. Slow transcription by RNA pol II can influence the folding of known RNA structures, such as the hairpin structure at the 3′ end of histone transcripts [129]. Overall, slow transcription speeds increase base-pairing in nascent RNAs and lead to more efficient splicing [12]. The speed of RNA polymerase II transcription is a function of modifications to its carboxy-terminal domain; these modifications can be influenced by many biological processes, reviewed in [130,131]. Cell signaling by the myc pathway influences RNA pol II elongation speed, suggesting that RNA structure and processing can be globally altered under certain conditions, reviewed in [130,132]. In addition to global changes, individual gene loci may be prone to fast or slow transcription speeds based on their DNA and chromatin composition. For example, DNA G-quadruplexes influence transcription speed [133]. Additionally, several DNA- and RNA-binding proteins influence RNA pol II modification and elongation speed [134]. Each of these factors can be regulated by other cellular pathways, potentially making transcription speed a dynamic factor regulated globally and fine-tuned at individual loci. More

Bioscience Reports (2023) **43** BSR20220149
https://doi.org/10.1042/BSR20220149

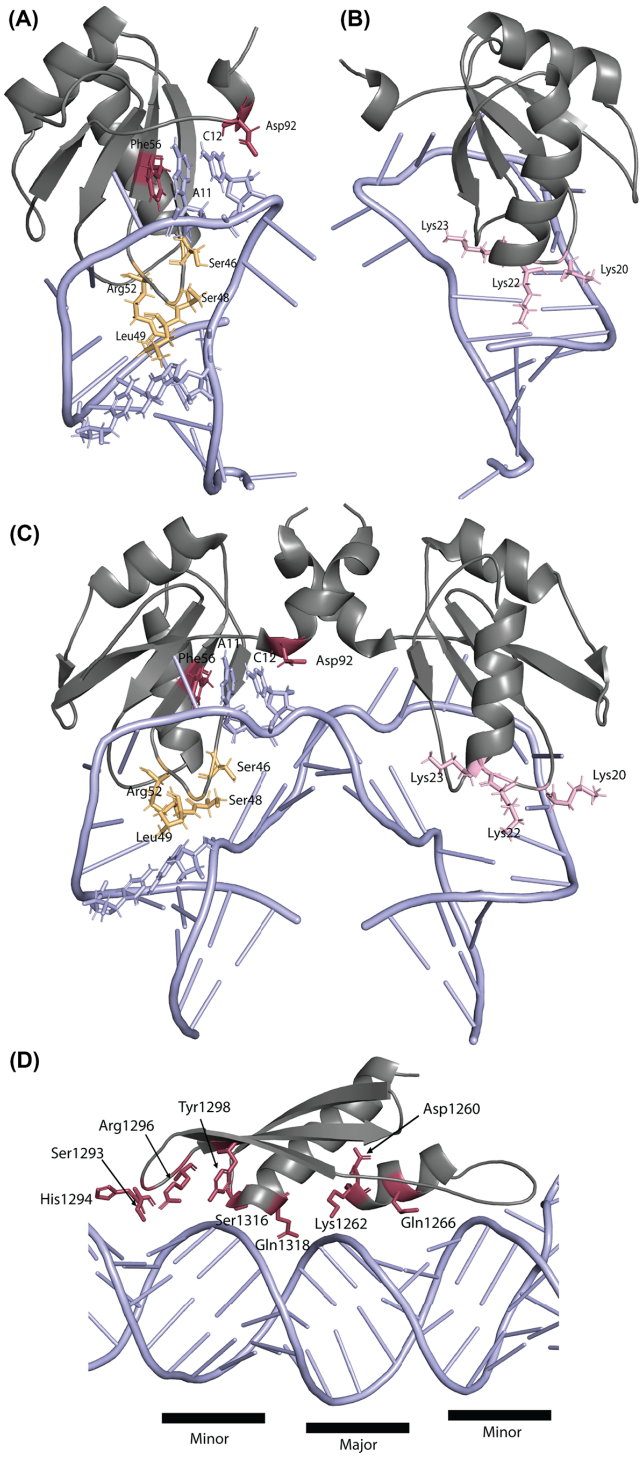

Figure 4. RNA interactions with RBPs

(**A**) The RNA-binding domain of U1A snRNP protein (gray) binds the U1 snRNA stem loop II (blue). Base stacking of A11 and C12 between amino acids Phe56 and Asp92 (maroon). Amino acids Ser46, Ser48, Leu49, and Arg52 (light orange) lock the protein into the hole defined by the RNA structure and interact with bases C11-G16. (PDB 1URN [115]). (**B**) U1A amino acids Lys20 and Lys22 contribute electrostatic interactions that stabilize the phosphodiester backbone of the RNA. Lys23 interacts with the U1A protein loop located in the open RNA (pink) (PDB 1URN [115]). (**C**) U1A RMM dimer binds PIE RNA structure. Amino acids are highlighted in the same colors as outlined above (PDB 1DZ5, [121]). (**D**) DROSHA dsRBD (PDB 6V5B [128]) in complex with pri-miR-16-2. Amino acids Ser1293, His1294, and Arg1296 interact with ribose in the minor groove and Tyr1298 interacts with the minor groove phosphate backbone. Gln1318 electrostatically interacts with the phosphate backbone of the major groove [127].

research is required to explain how perturbation of transcription speed influences the ensemble of structures formed by a nascent transcript and how a structural ensemble may influence processing of the precursor transcript and its downstream output.

Nucleotide modifications are another cellular mechanism that can impact RNA structure. There are about 180 known types of RNA nucleotide modifications [135]. Two common modifications are methyladenosine (m6A) and pseudouridine (pseudoU), reviewed by [136,137]. Methylation of adenosine at the 6th position makes the residue more likely to be unpaired, resulting in changes to RNA structure than can affect RBP interactions [138,139]. Modification of uridine to pseudoU stiffens the RNA backbone, usually resulting in more stable RNA structure, but the impact on structure is dependent on the context of the pseudoU modification [140]. Lack of pseudoU modification is associated with altered structural dynamics in the ribosome, specifically in the way the sections of the ribosome rotate with respect to one another [141]. The ability of pseudoU to change the structure of the ribosome and modulate its dynamics probably contributes to the translational defect in cells with ribosomes that lack pseudoU [141]. Although modifications are the rule for noncoding RNAs such as ribosomal RNA and tRNAs, RNA modifications are also present in pre-processed mRNAs, reviewed in [137]. For example, both m6A and pseudoU are common in precursor RNAs and affect splicing [142–144]. Updates to computational modeling algorithms are beginning to consider the effect of m6A on the biophysical characteristics of structure folding [24]. The growing volume of functional RNA modifications suggests that these modifications constitute a dynamic cellular mechanism that regulates RNA structure.

RNA-binding helicases can change RNA structures in the cell. At least eight helicases, including DEAH-box helicases DHX19, DHX38, DHX8 and DHX15, are essential for splicing in human cells, reviewed in [145,146]. These helicases primarily assist with the release of splicing factors and precursor RNA at different stages of the splicing cycle. In addition, several helicases recognize stalled or improper splicing and are associated with degradation of these precursor RNAs [145]. In precursor microRNAs (pri-miRNAs), the DEAD-box helicase DDX3X impacts the structural flexibility of the pri-miRNA and influences alternative processing by Drosha [123]. DDX3X binds double-stranded RNA as a dimer, with one DDX3X primarily interacting with one RNA strand [147]. DDX3X regulation of pri-miRNAs results in differences in mature miRNA isoform composition across tissues and between normal and cancerous samples [123]. There are at least 64 human RNA helicases, primarily identified by homology to DEAD and DEAH helicases [148]. These helicases and other RBPs can alter RNA structure by mechanisms other than conventional helicase unwinding, reviewed in [112]. The interaction between helicases and RNAs is often dependent on the modification state of the helicase, allowing for dynamic regulation of RNA:protein interaction [149]. Additional research is needed to understand the function and regulation of helicases.

There is evidence that RNA structure is altered at different stages of an RNA molecule's life cycle. Liu et al. found major differences between nuclear and cytoplasmic RNA structures in *Arabidopsis*, which suggested that RNA structure changed significantly from the precursor RNA to the mature transcript [10]. Structures in 3′UTRs can vary during different stages of development in zebrafish [11]. Differences in structure between RNAs *in vivo* versus purified from cells have been documented in yeast [4]. These studies demonstrate that RNA structure is not static. It is not clear in humans how structure is altered during processing of particular transcripts and whether refolding is a general characteristic that alters structure in predictable ways. Because RNA structure guides interactions with regulatory RBPs and nucleic acids, refolding of transcripts during processing has downstream implications for the fate of the transcript and gene expression.

What is the influence of precursor RNA structure on gene expression?

In this review we have discussed multiple examples of how RNA structures in human precursor RNAs impact RNA processing and alluded to the effect of altered processing on subsequent gene expression. One impact of altered splicing is production of an alternative protein isoform (Figure 5A–C). In *MAPT*, when exon 10 is skipped the mature transcript produces a Tau protein with three rather than four microtubule binding domains (Figure 5A) [48]. These Tau protein isoforms have different biological activities and altering their ratio is correlated with development of frontotemporal dementia and other neurodegenerative diseases [48,49]. In *SMN2, e*xon 7 is normally skipped, resulting in an unstable SMN protein isoform (Figure 5B) [150]. Spinal muscle atrophy occurs when neither *SMN1* nor *SMN2* can produce stable SMN protein. Exon 5 skipping in *MBNL1* RNA results in loss of part of the bipartite nuclear localization element in MBNL1 protein and cell-wide localization rather than nuclear localization [151–153] (Figure 5C). The sequestration of MBNL1 in toxic repeats is an important factor in myotonic dystrophy [154]. RNA degradation can also be influenced by RNA structures that affect processing. For example, *U1A* precursor RNAs contains a PIE element that is bound by two U1A proteins [118]. Binding of U1A to its own transcript results in inhibition

Bioscience Reports (2023) **43** BSR20220149
https://doi.org/10.1042/BSR20220149

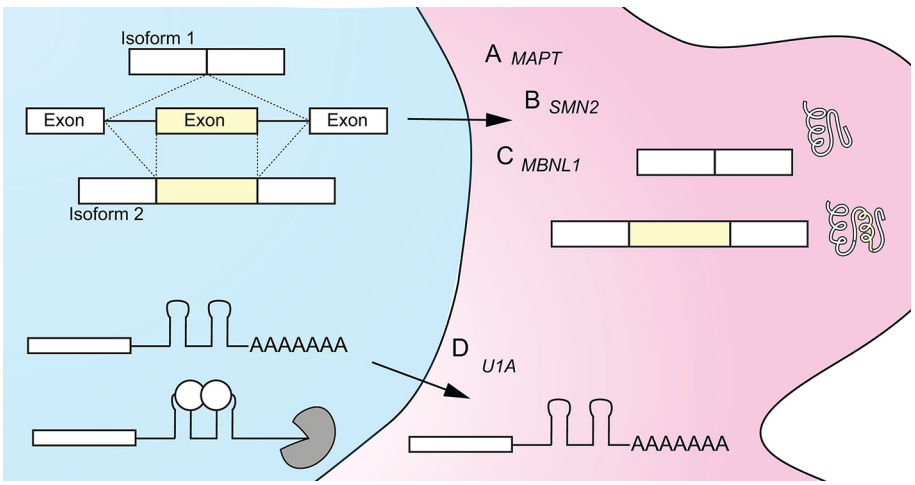

Figure 5. RNA structure impacts precursor RNA processing and influences gene expression

Schematic showing RNA processing in the nucleus (left, blue) and the impact of processing on gene expression in the cytoplasm (right, pink). (**A**) Alternative splicing can result in either exon skipping or exon inclusion (left, top). In *MAPT* RNA, exon skipping is promoted by hairpin formation at the 5′ splice site of exon 10. Exon skipping produces 3R and 4R transcripts and their corresponding protein isoforms. The 3R and 4R MAPT proteins have different biological functions. (**B**) MBNL1 binding to *MBNL1* RNA at the 3′ splice site of intron 4 promotes exon skipping. Exon skipping produces a transcript missing a bipartite nuclear localization motif and a cell-wide protein isoform. Exon inclusion produces a transcript that is translated into a nuclear MBNL1 protein isoform. (**C**) In *SMN2* RNA, exon skipping is promoted by hairpin formation at the 5′ splice site of exon 7. Exon skipping results in a protein isoform of SMN that is less stable than the full length SMN containing exon 7. Levels of SMN are associated with the severity of spinal muscular atrophy. (**D**) Most RNAs are polyadenylated at the 3′UTR (bottom). In *SMN2* processing the 3′ end of the transcript contains a PIE structural element bound by U1A that inhibits polyadenylation. Little or no polyadenylation leads to transcript instability and a decrease in RNA levels.

of polyadenylation and a decrease in *U1A* RNA [118,155] (Figure 5D). PIE elements in *SMN2* and other transcripts also can be bound by U1A and inhibit polyadenylation, ultimately resulting in lower levels of RNA [118,119]. Altered RNA processing can also influence nonsense-mediated decay [156], RNA localization [157] and protein expression [158]. Nascent RNA structure has ripple effects on all aspects of the RNA life cycle and can contribute to human diseases.

How can precursor RNA structure be targeted by therapeutics?

Because RNA structure influences its functional interactions with other molecules, structure is a target for intervention, including at the nascent RNA stage. Antisense oligonucleotides (ASOs) can be designed to alter RNA structure, reviewed in [159,160]. In a structured RNA, bases normally interact in *cis* to form standard hairpins or stem-loop structures. An ASO can compete for base pairing in *trans*. The hybridization between the ASO and its target RNA opens up nucleobases for interaction with other nucleotides or proteins and could have global effects on structure. In the 7SK snRNP structural rearrangement is important for release of kinases involved in phosphorylation of the Poll II carboxy-terminal domain, leading to transcriptional control. ASOs that target sequences within 7SK dynamic hairpins block the structural transition of 7SK from one state to another and alter the ability of 7SK to regulate transcription [161]. The SARS-CoV2 corona virus has a highly structured single-stranded RNA genome [162]. One strategy currently under development for treatment of SARS infection is an ASO designed to disrupt a 3′ stem-loop involved in viral replication, reviewed in [163]. A similar mechanism of competing hybridization has been used to develop toehold switches, which switch structures in the presence of a particular RNA sequence to allow translation [164]. Toehold sensors have been developed for many applications including as a method to detect viral infections like SARS-CoV2 and Zika [165,166] and identify genomic variation [167]. There is software available to design toehold structures (NUPACK) [168]. Other hybridization methods that impact RNA structure have been developed to target transcriptional regulation [169].

Although hybridization offers a straightforward mechanism of structure change, ASOs are difficult to deliver to human tissue, whereas small molecules are generally more tractable for medical treatment. Small molecules can target and stabilize or destabilize specific RNA structures, reviewed in [170–172]. The capacity of small molecules to act on RNA structures was evident early on from bacterial riboswitches, which are designed to recognize a variety of different small molecules (e.g., metabolites) and change conformation to effect transcriptional or translational regulation, reviewed in [173]. Although most proof-of-principle molecules target noncoding or viral RNAs [174,175], small molecules that target RNA structures can be used to control nascent RNA processing. The FDA approved small molecule risdiplam has been developed to target *SMN2* splicing at exon 7, reviewed in [176]. Although the exact mechanism of action is still under investigation, these molecules may function to stabilize the interaction between the U1 spliceosome and the 5′ splice site [177,178]. Modulating RNA structure to diagnosis or to treat disease is a rapidly growing field. Targeting function RNA structures in precursor RNAs is an important direction for therapeutic development.

Summary

RNA molecules are naturally structured. Due to the low abundance, long-length, and flexible nature of nascent RNAs, precursor RNA structure is understudied. New structural methods are continuing to advance our technological capabilities and document structures within precursor RNAs. In particular, chemical probing and cryo-EM methods have expanded our understanding of secondary and tertiary structures. However, even when structural models are available, it is difficult to identify functional structures and understand their mechanisms. Structures within precursor RNAs determine how nascent transcripts interact with protein and nucleic acid co-factors. By influencing these interactions, RNA structure influences RNA processing pathways. Most studies have focused on the impact of structure on splicing and polyadenylation, but future research may tell us more about how RNA structure affects other processing pathways like RNA editing. Due to their impact on processing, RNA structures impact gene expression and play a role in disease. We are beginning to develop antisense oligonucleotides and small molecule methods to alter RNA structure *in vivo*; these methods can be broadly applied to target functional RNA structures, including those that regulate RNA processing.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

This work was supported by the National Institutes of Health [grant number R35GM142851].

CRediT Author Contribution

Austin Herbert: Conceptualization, Formal analysis, Investigation, Visualization, Writing—review & editing. **Abigail Hatfield:** Conceptualization, Writing—review & editing. **Lela Lackey:** Conceptualization, Visualization, Writing—original draft, Writing—review & editing.

Acknowledgements

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from: GTEx Analysis Release V8 (dbGaP Accession phs000424.v8.p2) from the GTEx Portal on 6/01/2022.

Abbreviations

ASO, antisense oligonucleotide; dsRBD, double-stranded RNA-binding domain; eCLIP, enhanced cross-linking and immunoprecipitation; RBD, RNA-binding domain; RNP, Ribonucleoprotein; RRM, RNA recognition motif.

References

1 Ganser, L.R., Kelly, M.L., Herschlag, D. and Al-Hashimi, H.M. (2019) The roles of structural dynamics in the cellular functions of RNAs. *Nat. Rev. Mol. Cell Biol.* **20**, 474–489, https://doi.org/10.1038/s41580-019-0136-0

2 Mustoe, A.M., Brooks, C.L. and Al-Hashimi, H.M. (2014) Hierarchy of RNA functional dynamics. *Annu. Rev. Biochem.* **83**, 441–466, https://doi.org/10.1146/annurev-biochem-060713-035524

3 Wan, Y., Qu, K., Zhang, Q.C., Flynn, R.A., Manor, O., Ouyang, Z. et al. (2014) Landscape and variation of RNA secondary structure across the human transcriptome. *Nature* **505**, 706–709, https://doi.org/10.1038/nature12946
4 Rouskin, S., Zubradt, M., Washietl, S., Kellis, M. and Weissman, J.S. (2014) Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* **505**, 701–705, https://doi.org/10.1038/nature12894
5 Ding, Y., Tang, Y., Kwok, C.K., Zhang, Y., Bevilacqua, P.C. and Assmann, S.M. (2014) In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **505**, 696–700, https://doi.org/10.1038/nature12756
6 Incarnato, D., Neri, F., Anselmi, F. and Oliviero, S. (2014) Genome-wide profiling of mouse RNA secondary structures reveals key features of the mammalian transcriptome. *Genome Biol.* **15**, 491, https://doi.org/10.1186/s13059-014-0491-2
7 Kertesz, M., Wan, Y., Mazor, E., Rinn, J.L., Nutter, R.C., Chang, H.Y. et al. (2010) Genome-wide measurement of RNA secondary structure in yeast. *Nature* **467**, 103–107, https://doi.org/10.1038/nature09322
8 Spitale, R.C., Flynn, R.A., Zhang, Q.C., Crisalli, P., Lee, B., Jung, J.W. et al. (2015) Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* **519**, 486–490, https://doi.org/10.1038/nature14263
9 Vicens, Q. and Kieft, J.S. (2022) Thoughts on how to think (and talk) about RNA structure. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2112677119, https://doi.org/10.1073/pnas.2112677119
10 Liu, Z., Liu, Q., Yang, X., Zhang, Y., Norris, M., Chen, X. et al. (2021) In vivo nuclear RNA structurome reveals RNA-structure regulation of mRNA processing in plants. *Genome Biol.* **22**, 11, https://doi.org/10.1186/s13059-020-02236-4
11 Shi, B., Zhang, J., Heng, J., Gong, J., Zhang, T., Li, P. et al. (2020) RNA structural dynamics regulate early embryogenesis through controlling transcriptome fate and function. *Genome Biol.* **21**, 120, https://doi.org/10.1186/s13059-020-02022-2
12 Saldi, T., Riemondy, K., Erickson, B. and Bentley, D.L. (2021) Alternative RNA structures formed during transcription depend on elongation rate and modify RNA processing. *Mol. Cell.* **81**, 1789e5–1801e5, https://doi.org/10.1016/j.molcel.2021.01.040
13 Wang, X.W., Liu, C.X., Chen, L.L. and Zhang, Q.C. (2021) RNA structure probing uncovers RNA structure-dependent biological functions. *Nat. Chem. Biol.* **17**, 755–766, https://doi.org/10.1038/s41589-021-00805-7
14 Zubradt, M., Gupta, P., Persad, S., Lambowitz, A.M., Weissman, J.S. and Rouskin, S. (2017) DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods* **14**, 75–82, https://doi.org/10.1038/nmeth.4057
15 Siegfried, N.A., Busan, S., Rice, G.M., Nelson, J.A. and Weeks, K.M. (2014) RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* **11**, 959–965, https://doi.org/10.1038/nmeth.3029
16 Busan, S., Weidmann, C.A., Sengupta, A. and Weeks, K.M. (2019) Guidelines for SHAPE reagent choice and detection strategy for RNA structure probing studies. *Biochemistry* **58**, 2655–2664, https://doi.org/10.1021/acs.biochem.8b01218
17 Wang, P.Y., Sexton, A.N., Culligan, W.J. and Simon, M.D. (2019) Carbodiimide reagents for the chemical probing of RNA structure in cells. *RNA* **25**, 135–146, https://doi.org/10.1261/rna.067561.118
18 Kastner, B., Will, C.L., Stark, H. and Luhrmann, R. (2019) Structural insights into nuclear pre-mRNA splicing in higher eukaryotes. *Cold Spring Harb. Perspect. Biol.* **11**, a032417, https://doi.org/10.1101/cshperspect.a032417
19 Earl, L.A., Falconieri, V., Milne, J.L. and Subramaniam, S. (2017) Cryo-EM: beyond the microscope. *Curr. Opin. Struct. Biol.* **46**, 71–78, https://doi.org/10.1016/j.sbi.2017.06.002
20 Solayman, M., Litfin, T., Singh, J., Paliwal, K., Zhou, Y. and Zhan, J. (2022) Probing RNA structures and functions by solvent accessibility: an overview from experimental and computational perspectives. *Brief. Bioinform.* **23**, bbac112, https://doi.org/10.1093/bib/bbac112
21 Sato, K., Akiyama, M. and Sakakibara, Y. (2021) RNA secondary structure prediction using deep learning with thermodynamic integration. *Nat. Commun.* **12**, 941, https://doi.org/10.1038/s41467-021-21194-4
22 Fu, L., Cao, Y., Wu, J., Peng, Q., Nie, Q. and Xie, X. (2022) UFold: fast and accurate RNA secondary structure prediction with deep learning. *Nucleic Acids Res.* **50**, e14, https://doi.org/10.1093/nar/gkab1074
23 Townshend, R.J.L., Eismann, S., Watkins, A.M., Rangan, R., Karelina, M., Das, R. et al. (2021) Geometric deep learning of RNA structure. *Science* **373**, 1047–1051, https://doi.org/10.1126/science.abe5650
24 Kierzek, E., Zhang, X., Watson, R.M., Kennedy, S.D., Szabat, M., Kierzek, R. et al. (2022) Secondary structure prediction for RNA sequences including N(6)-methyladenosine. *Nat. Commun.* **13**, 1271, https://doi.org/10.1038/s41467-022-28817-4
25 Reuter, J.S. and Mathews, D.H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* **11**, 129, https://doi.org/10.1186/1471-2105-11-129
26 Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415, https://doi.org/10.1093/nar/gkg595
27 Zhang, H., Zhang, L., Mathews, D.H. and Huang, L. (2020) LinearPartition: linear-time approximation of RNA folding partition function and base-pairing probabilities. *Bioinformatics* **36**, i258–i267, https://doi.org/10.1093/bioinformatics/btaa460
28 Szikszai, M., Wise, M., Datta, A., Ward, M. and Mathews, D.H. (2022) Deep learning models for RNA secondary structure prediction (probably) do not generalise across families. *Bioinformatics* 3892–3899, https://doi.org/10.1093/bioinformatics/btac415
29 Li, B., Cao, Y., Westhof, E. and Miao, Z. (2020) Advances in RNA 3D Structure modeling using experimental data. *Front. Genet.* **11**, 574485, https://doi.org/10.3389/fgene.2020.574485
30 Mathews, D.H. (2019) How to benchmark RNA secondary structure prediction accuracy. *Methods* **162-163**, 60–67, https://doi.org/10.1016/j.ymeth.2019.04.003
31 Bubenik, J.L., Hale, M., McConnell, O., Wang, E., Swanson, M.S., Spitale, R. et al. (2020) RNA structure probing to characterize RNA-protein interactions on a low abundance pre-mRNA in living cells. *RNA* 343–358
32 Kwok, C.K., Ding, Y., Tang, Y., Assmann, S.M. and Bevilacqua, P.C. (2013) Determination of in vivo RNA structure in low-abundance transcripts. *Nat. Commun.* **4**, 2971, https://doi.org/10.1038/ncomms3971

33 Smola, M.J., Calabrese, J.M. and Weeks, K.M. (2015) Detection of RNA-protein interactions in living cells with SHAPE. *Biochemistry* **54**, 6867–6875, https://doi.org/10.1021/acs.biochem.5b00977
34 Flynn, R.A., Zhang, Q.C., Spitale, R.C., Lee, B., Mumbach, M.R. and Chang, H.Y. (2016) Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. *Nat. Protoc.* **11**, 273–290, https://doi.org/10.1038/nprot.2016.011
35 Tomezsko, P.J., Corbin, V.D.A., Gupta, P., Swaminathan, H., Glasgow, M., Persad, S. et al. (2020) Determination of RNA structural diversity and its role in HIV-1 RNA splicing. *Nature* **582**, 438–442, https://doi.org/10.1038/s41586-020-2253-5
36 Woods, C.T., Lackey, L., Williams, B., Dokholyan, N.V., Gotz, D. and Laederach, A. (2017) Comparative visualization of the RNA suboptimal conformational ensemble in vivo. *Biophys. J.* **113**, 290–301, https://doi.org/10.1016/j.bpj.2017.05.031
37 Aviran, S. and Incarnato, D. (2022) Computational approaches for RNA structure ensemble deconvolution from structure probing data. *J. Mol. Biol.* **434**, 167635, https://doi.org/10.1016/j.jmb.2022.167635
38 Kalmykova, S., Kalinina, M., Denisov, S., Mironov, A., Skvortsov, D., Guigo, R. et al. (2021) Conserved long-range base pairings are associated with pre-mRNA processing of human genes. *Nat. Commun.* **12**, 2300, https://doi.org/10.1038/s41467-021-22549-7
39 Ermolenko, D.N. and Mathews, D.H. (2021) Making ends meet: new functions of mRNA secondary structure. *Wiley Interdiscip .Rev. RNA* **12**, e1611, https://doi.org/10.1002/wrna.1611
40 Kessler, O., Jiang, Y. and Chasin, L.A. (1993) Order of intron removal during splicing of endogenous adenine phosphoribosyltransferase and dihydrofolate reductase pre-mRNA. *Mol. Cell. Biol.* **13**, 6211–6222
41 Kim, S.W., Taggart, A.J., Heintzelman, C., Cygan, K.J., Hull, C.G., Wang, J. et al. (2017) Widespread intra-dependencies in the removal of introns from human transcripts. *Nucleic Acids Res.* **45**, 9503–9513, https://doi.org/10.1093/nar/gkx661
42 Drexler, H.L., Choquet, K. and Churchman, L.S. (2020) Splicing kinetics and coordination revealed by direct nascent RNA sequencing through nanopores. *Mol. Cell.* **77**, 985e8–998e8, https://doi.org/10.1016/j.molcel.2019.11.017
43 Yu, A.M., Gasper, P.M., Cheng, L., Lai, L.B., Kaur, S., Gopalan, V. et al. (2021) Computationally reconstructing cotranscriptional RNA folding from experimental data reveals rearrangement of non-native folding intermediates. *Mol. Cell.* **81**, 870e10–883e10, https://doi.org/10.1016/j.molcel.2020.12.017
44 Eperon, L.P., Estibeiro, J.P. and Eperon, I.C. (1986) The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA. *Nature* **324**, 280–282, https://doi.org/10.1038/324280a0
45 Eperon, L.P., Graham, I.R., Griffiths, A.D. and Eperon, I.C. (1988) Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell* **54**, 393–401, https://doi.org/10.1016/0092-8674(88)90202-4
46 Roca, X., Krainer, A.R. and Eperon, I.C. (2013) Pick one, but be quick: 5′ splice sites and the problems of too many choices. *Genes Dev.* **27**, 129–144, https://doi.org/10.1101/gad.209759.112
47 Park, S.A., Ahn, S.I. and Gallo, J.M. (2016) Tau mis-splicing in the pathogenesis of neurodegenerative disorders. *BMB Rep.* **49**, 405–413, https://doi.org/10.5483/BMBRep.2016.49.8.084
48 Hefti, M.M., Farrell, K., Kim, S., Bowles, K.R., Fowkes, M.E., Raj, T. et al. (2018) High-resolution temporal and regional mapping of MAPT expression and splicing in human brain development. *PloS ONE* **13**, e0195771, https://doi.org/10.1371/journal.pone.0195771
49 Hutton, M., Lendon, C.L., Rizzu, P., Baker, M., Froelich, S., Houlden, H. et al. (1998) Association of missense and 5′-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* **393**, 702–705, https://doi.org/10.1038/31508
50 Lisowiec, J., Magner, D., Kierzek, E., Lenartowicz, E. and Kierzek, R. (2015) Structural determinants for alternative splicing regulation of the MAPT pre-mRNA. *RNA Biol.* **12**, 330–342, https://doi.org/10.1080/15476286.2015.1017214
51 Kumar, J., Lackey, L., Waldern, J.M., Dey, A., Mustoe, A.M., Weeks, K. et al. (2022) Quantitative prediction of variant effects on alternative splicing in MAPT using endogenous pre-messenger RNA structure probing. *Elife* **11**, e73888, https://doi.org/10.7554/eLife.73888
52 Weldon, C., Behm-Ansmant, I., Hurley, L.H., Burley, G.A., Branlant, C., Eperon, I.C. et al. (2017) Identification of G-quadruplexes in long functional RNAs using 7-deazaguanine RNA. *Nat. Chem. Biol.* **13**, 18–20, https://doi.org/10.1038/nchembio.2228
53 Yadegari, H., Biswas, A., Akhter, M.S., Driesen, J., Ivaskevicius, V., Marquardt, N. et al. (2016) Intron retention resulting from a silent mutation in the VWF gene that structurally influences the 5′ splice site. *Blood* **128**, 2144–2152, https://doi.org/10.1182/blood-2016-02-699686
54 Singh, N.N., Luo, D. and Singh, R.N. (2018) Pre-mRNA splicing modulation by antisense oligonucleotides. *Methods Mol. Biol.* **1828**, 415–437, https://doi.org/10.1007/978-1-4939-8651-4˙26
55 Singh, N.N., Singh, R.N. and Androphy, E.J. (2007) Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes. *Nucleic Acids Res.* **35**, 371–389, https://doi.org/10.1093/nar/gkl1050
56 Buratti, E., Muro, A.F., Giombi, M., Gherbassi, D., Iaconcig, A. and Baralle, F.E. (2004) RNA folding affects the recruitment of SR proteins by mouse and human polypurinic enhancer elements in the fibronectin EDA exon. *Mol. Cell. Biol.* **24**, 1387–1400, https://doi.org/10.1128/MCB.24.3.1387-1400.2004
57 Saha, K., Fernandez, M.M., Biswas, T., Joseph, S. and Ghosh, G. (2021) Discovery of a pre-mRNA structural scaffold as a contributor to the mammalian splicing code. *Nucleic Acids Res.* **49**, 7103–7121, https://doi.org/10.1093/nar/gkab533
58 Saha, K. and Ghosh, G. (2022) Cooperative engagement and subsequent selective displacement of SR proteins define the pre-mRNA 3D structural scaffold for early spliceosome assembly. *Nucleic Acids Res.* **50**, 8262–8278, https://doi.org/10.1093/nar/gkac636
59 Gahura, O., Hammann, C., Valentova, A., Puta, F. and Folk, P. (2011) Secondary structure is required for 3′ splice site recognition in yeast. *Nucleic Acids Res.* **39**, 9759–9767, https://doi.org/10.1093/nar/gkr662
60 Wu, X. and Bartel, D.P. (2017) Widespread influence of 3′-end structures on mammalian mRNA processing and stability. *Cell* **169**, 905e11–917e11, https://doi.org/10.1016/j.cell.2017.04.036
61 Fish, L., Khoroshkin, M., Navickas, A., Garcia, K., Culbertson, B., Hanisch, B. et al. (2021) A prometastatic splicing program regulated by SNRPA1 interactions with structured RNA elements. *Science* **372**, eabc7531, https://doi.org/10.1126/science.abc7531

62 Gates, D.P., Coonrod, L.A. and Berglund, J.A. (2011) Autoregulated splicing of muscleblind-like 1 (MBNL1) Pre-mRNA. *J. Biol. Chem.* **286**, 34224–34233, https://doi.org/10.1074/jbc.M111.236547
63 Warf, M.B., Diegel, J.V., von Hippel, P.H. and Berglund, J.A. (2009) The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9203–9208, https://doi.org/10.1073/pnas.0900342106
64 Gosai, S.J., Foley, S.W., Wang, D., Silverman, I.M., Selamoglu, N., Nelson, A.D. et al. (2015) Global analysis of the RNA-protein interaction and RNA secondary structure landscapes of the Arabidopsis nucleus. *Mol. Cell.* **57**, 376–388, https://doi.org/10.1016/j.molcel.2014.12.004
65 Saha, K., England, W., Fernandez, M.M., Biswas, T., Spitale, R.C. and Ghosh, G. (2020) Structural disruption of exonic stem-loops immediately upstream of the intron regulates mammalian splicing. *Nucleic Acids Res.* **48**, 6294–6309, https://doi.org/10.1093/nar/gkaa358
66 Sun, L., Fazal, F.M., Li, P., Broughton, J.P., Lee, B., Tang, L. et al. (2019) RNA structure maps across mammalian cellular compartments. *Nat. Struct. Mol. Biol.* **26**, 322–330, https://doi.org/10.1038/s41594-019-0200-7
67 Zafrir, Z. and Tuller, T. (2015) Nucleotide sequence composition adjacent to intronic splice sites improves splicing efficiency via its effect on pre-mRNA local folding in fungi. *RNA* **21**, 1704–1718, https://doi.org/10.1261/rna.051268.115
68 Warf, M.B. and Berglund, J.A. (2010) Role of RNA structure in regulating pre-mRNA splicing. *Trends Biochem. Sci.* **35**, 169–178, https://doi.org/10.1016/j.tibs.2009.10.004
69 Xu, B., Meng, Y. and Jin, Y. (2021) RNA structures in alternative splicing and back-splicing. *Wiley Interdiscip. Rev. RNA* **12**, e1626, https://doi.org/10.1002/wrna.1626
70 Singh, N.N., Androphy, E.J. and Singh, R.N. (2004) In vivo selection reveals combinatorial controls that define a critical exon in the spinal muscular atrophy genes. *RNA* **10**, 1291–1305, https://doi.org/10.1261/rna.7580704
71 Singh, N.N., Lawler, M.N., Ottesen, E.W., Upreti, D., Kaczynski, J.R. and Singh, R.N. (2013) An intronic structure enabled by a long-distance interaction serves as a novel target for splicing correction in spinal muscular atrophy. *Nucleic Acids Res.* **41**, 8144–8165, https://doi.org/10.1093/nar/gkt609
72 Singh, N.N., Lee, B.M., DiDonato, C.J. and Singh, R.N. (2015) Mechanistic principles of antisense targets for the treatment of spinal muscular atrophy. *Future Med. Chem.* **7**, 1793–1808, https://doi.org/10.4155/fmc.15.101
73 Singh, N.N., O'Leary, C.A., Eich, T., Moss, W.N. and Singh, R.N. (2022) Structural Context of a Critical Exon of Spinal Muscular Atrophy Gene. *Front Mol. Biosci.* **9**, 928581, https://doi.org/10.3389/fmolb.2022.928581
74 Miyaso, H., Okumura, M., Kondo, S., Higashide, S., Miyajima, H. and Imaizumi, K. (2003) An intronic splicing enhancer element in survival motor neuron (SMN) Pre-mRNA. *J. Biol. Chem.* **278**, 15825–15831, https://doi.org/10.1074/jbc.M209271200
75 Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L. et al. (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **80**, 155–165, https://doi.org/10.1016/0092-8674(95)90460-3
76 Monani, U.R., Lorson, C.L., Parsons, D.W., Prior, T.W., Androphy, E.J., Burghes, A.H. et al. (1999) A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. *Hum. Mol. Genet.* **8**, 1177–1183, https://doi.org/10.1093/hmg/8.7.1177
77 Cui, J. and Placzek, W.J. (2018) Post-transcriptional regulation of anti-apoptotic BCL2 family members. *Int. J. Mol. Sci.* **19**, 308, https://doi.org/10.3390/ijms19010308
78 Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D. et al. (2004) The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **32**, D493–D496, https://doi.org/10.1093/nar/gkh103
79 Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K. et al. (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050, https://doi.org/10.1101/gr.3715005
80 Zeng, Y., Zeng, H., Fair, B.J., Krishnamohan, A., Hou, Y., Hall, J.M. et al. (2022) Profiling lariat intermediates reveals genetic determinants of early and late co-transcriptional splicing. *Mol. Cell.* **82**, 4681–99e8, https://doi.org/10.1016/j.molcel.2022.11.004
81 Eddy, S.R. and Durbin, R. (1994) RNA sequence analysis using covariance models. *Nucleic Acids Res.* **22**, 2079–2088, https://doi.org/10.1093/nar/22.11.2079
82 Zhang, J., Fei, Y., Sun, L. and Zhang, Q.C. (2022) Advances and opportunities in RNA structure experimental determination and computational modeling. *Nat. Methods* **19**, 1193–1207, https://doi.org/10.1038/s41592-022-01623-y
83 Rivas, E., Clements, J. and Eddy, S.R. (2017) A statistical test for conserved RNA structure shows lack of evidence for structure in lncRNAs. *Nat. Methods* **14**, 45–48, https://doi.org/10.1038/nmeth.4066
84 Rivas, E. (2020) RNA structure prediction using positive and negative evolutionary information. *PLoS Comput. Biol.* **16**, e1008387, https://doi.org/10.1371/journal.pcbi.1008387
85 Consortium, E.P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74, https://doi.org/10.1038/nature11247
86 Luo, Y., Hitz, B.C., Gabdank, I., Hilton, J.A., Kagda, M.S., Lam, B. et al. (2020) New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. *Nucleic Acids Res.* **48**, D882–D889, https://doi.org/10.1093/nar/gkz1062
87 Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B. et al. (2016) Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514, https://doi.org/10.1038/nmeth.3810
88 Gautrey, H.L. and Tyson-Capper, A.J. (2012) Regulation of Mcl-1 by SRSF1 and SRSF5 in cancer cells. *PLoS ONE* **7**, e51497, https://doi.org/10.1371/journal.pone.0051497
89 Tyson-Capper, A. and Gautrey, H. (2018) Regulation of Mcl-1 alternative splicing by hnRNP F, H1 and K in breast cancer cells. *RNA Biol.* **15**, 1448–1457, https://doi.org/10.1080/15476286.2018.1551692
90 Singh, R.N. and Singh, N.N. (2018) Mechanism of splicing regulation of spinal muscular atrophy genes. *Adv. Neurobiol.* **20**, 31–61, https://doi.org/10.1007/978-3-319-89689-2˙2
91 Singh, N.N., Seo, J., Ottesen, E.W., Shishimorova, M., Bhattacharya, D. and Singh, R.N. (2011) TIA1 prevents skipping of a critical exon associated with spinal muscular atrophy. *Mol. Cell. Biol.* **31**, 935–954, https://doi.org/10.1128/MCB.00945-10

92 Sun, L., Xu, K., Huang, W., Yang, Y.T., Li, P., Tang, L. et al. (2021) Predicting dynamic cellular protein-RNA interactions by deep learning using in vivo RNA structures. *Cell Res.* **31**, 495–516, https://doi.org/10.1038/s41422-021-00476-y
93 Landrum, M.J., Lee, J.M., Benson, M., Brown, G.R., Chao, C., Chitipiralla, S. et al. (2018) ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **46**, D1062–D1067, https://doi.org/10.1093/nar/gkx1153
94 Stenson, P.D., Ball, E.V., Mort, M., Phillips, A.D., Shiel, J.A., Thomas, N.S. et al. (2003) Human Gene Mutation Database (HGMD): 2003 update. *Hum. Mutat.* **21**, 577–581, https://doi.org/10.1002/humu.10212
95 Consortium, G.T. (2020) The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330, https://doi.org/10.1126/science.aaz1776
96 Garrido-Martin, D., Borsari, B., Calvo, M., Reverter, F. and Guigo, R. (2021) Identification and analysis of splicing quantitative trait loci across multiple tissues in the human genome. *Nat. Commun.* **12**, 727, https://doi.org/10.1038/s41467-020-20578-2
97 Mittleman, B.E., Pott, S., Warland, S., Zeng, T., Mu, Z., Kaur, M. et al. (2020) Alternative polyadenylation mediates genetic regulation of gene expression. *Elife* **9**, e57492, https://doi.org/10.7554/eLife.57492
98 Tate, J.G., Bamford, S., Jubb, H.C., Sondka, Z., Beare, D.M., Bindal, N. et al. (2019) COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.* **47**, D941–D947, https://doi.org/10.1093/nar/gky1015
99 Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alfoldi, J., Wang, Q. et al. (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443, https://doi.org/10.1038/s41586-020-2308-7
100 Halvorsen, M., Martin, J.S., Broadaway, S. and Laederach, A. (2010) Disease-associated mutations that alter the RNA structural ensemble. *PLos Genet.* **6**, e1001074, https://doi.org/10.1371/journal.pgen.1001074
101 Lin, J., Chen, Y., Zhang, Y. and Ouyang, Z. (2020) Identification and analysis of RNA structural disruptions induced by single nucleotide variants using Riprap and RiboSNitchDB. *NAR Genom. Bioinform.* **2**, lqaa057, https://doi.org/10.1093/nargab/lqaa057
102 Sabarinathan, R., Tafer, H., Seemann, S.E., Hofacker, I.L., Stadler, P.F. and Gorodkin, J. (2013) The RNAsnp web server: predicting SNP effects on local RNA secondary structure. *Nucleic Acids Res.* **41**, W475–W479, https://doi.org/10.1093/nar/gkt291
103 Sabarinathan, R., Tafer, H., Seemann, S.E., Hofacker, I.L., Stadler, P.F. and Gorodkin, J. (2013) RNAsnp: efficient detection of local RNA secondary structure changes induced by SNPs. *Hum. Mutat.* **34**, 546–556, https://doi.org/10.1002/humu.22273
104 Waldern, J.M., Kumar, J. and Laederach, A. (2022) Disease-associated human genetic variation through the lens of precursor and mature RNA structure. *Hum. Genet.* **141**, 1659–1672, https://doi.org/10.1007/s00439-021-02395-9
105 Lackey, L., Coria, A., Woods, C., McArthur, E. and Laederach, A. (2018) Allele-specific SHAPE-MaP assessment of the effects of somatic variation and protein binding on mRNA structure. *RNA* **24**, 513–528, https://doi.org/10.1261/rna.064469.117
106 Solem, A.C., Halvorsen, M., Ramos, S.B. and Laederach, A. (2015) The potential of the riboSNitch in personalized medicine. *Wiley Interdiscip. Rev. RNA* **6**, 517–532, https://doi.org/10.1002/wrna.1291
107 Bartys, N., Kierzek, R. and Lisowiec-Wachnicka, J. (2019) The regulation properties of RNA secondary structure in alternative splicing. *Biochim. Biophys. Acta Gene Regul. Mech.* **1862**, 194401, https://doi.org/10.1016/j.bbagrm.2019.07.002
108 Dominguez, D., Freese, P., Alexis, M.S., Su, A., Hochman, M., Palden, T. et al. (2018) Sequence, structure, and context preferences of human RNA binding proteins. *Mol. Cell.* **70**, 854e9–867e9, https://doi.org/10.1016/j.molcel.2018.05.001
109 Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R. et al. (2020) A large-scale binding and functional map of human RNA-binding proteins. *Nature* **583**, 711–719, https://doi.org/10.1038/s41586-020-2077-3
110 Van Nostrand, E.L., Pratt, G.A., Yee, B.A., Wheeler, E.C., Blue, S.M., Mueller, J. et al. (2020) Principles of RNA processing from analysis of enhanced CLIP maps for 150 RNA binding proteins. *Genome Biol.* **21**, 90, https://doi.org/10.1186/s13059-020-01982-9
111 Lambert, N., Robertson, A., Jangi, M., McGeary, S., Sharp, P.A. and Burge, C.B. (2014) RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol. Cell.* **54**, 887–900, https://doi.org/10.1016/j.molcel.2014.04.016
112 Hentze, M.W., Castello, A., Schwarzl, T. and Preiss, T. (2018) A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* **19**, 327–341, https://doi.org/10.1038/nrm.2017.130
113 Corley, M., Burns, M.C. and Yeo, G.W. (2020) How RNA-binding proteins interact with RNA: molecules and mechanisms. *Mol. Cell.* **78**, 9–29, https://doi.org/10.1016/j.molcel.2020.03.011
114 DeKoster, G.T., Delaney, K.J. and Hall, K.B. (2014) A compare-and-contrast NMR dynamics study of two related RRMs: U1A and SNF. *Biophys. J.* **107**, 208–219, https://doi.org/10.1016/j.bpj.2014.05.026
115 Oubridge, C., Ito, N., Evans, P.R., Teo, C.H. and Nagai, K. (1994) Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372**, 432–438, https://doi.org/10.1038/372432a0
116 Allain, F.H., Howe, P.W., Neuhaus, D. and Varani, G. (1997) Structural basis of the RNA-binding specificity of human U1A protein. *EMBO J.* **16**, 5764–5772, https://doi.org/10.1093/emboj/16.18.5764
117 Rimmele, M.E. and Belasco, J.G. (1998) Target discrimination by RNA-binding proteins: role of the ancillary protein U2A′ and a critical leucine residue in differentiating the RNA-binding specificity of spliceosomal proteins U1A and U2B″. *RNA* **4**, 1386–1396, https://doi.org/10.1017/S1355838298981171
118 Boelens, W.C., Jansen, E.J., van Venrooij, W.J., Stripecke, R., Mattaj, I.W. and Gunderson, S.I. (1993) The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA. *Cell* **72**, 881–892, https://doi.org/10.1016/0092-8674(93)90577-D
119 Workman, E., Veith, A. and Battle, D.J. (2014) U1A regulates 3′ processing of the survival motor neuron mRNA. *J. Biol. Chem.* **289**, 3703–3712, https://doi.org/10.1074/jbc.M113.538264
120 Clerte, C. and Hall, K.B. (2004) Global and local dynamics of the U1A polyadenylation inhibition element (PIE) RNA and PIE RNA-U1A complexes. *Biochemistry* **43**, 13404–13415, https://doi.org/10.1021/bi049117g

121 Varani, L., Gunderson, S.I., Mattaj, I.W., Kay, L.E., Neuhaus, D. and Varani, G. (2000) The NMR structure of the 38 kDa U1A protein - PIE RNA complex reveals the basis of cooperativity in regulation of polyadenylation by human U1A protein. *Nat. Struct. Biol.* **7**, 329–335, https://doi.org/10.1038/74101
122 Gunderson, S.I., Beyer, K., Martin, G., Keller, W., Boelens, W.C. and Mattaj, L.W. (1994) The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell* **76**, 531–541, https://doi.org/10.1016/0092-8674(94)90116-3
123 Bofill-De Ros, X., Hong, Z., Birkenfeld, B., Alamo-Ortiz, S., Yang, A., Dai, L. et al. (2022) Flexible pri-miRNA structures enable tunable production of 5′ isomiRs. *RNA Biol.* **19**, 279–289, https://doi.org/10.1080/15476286.2022.2025680
124 Jouravleva, K., Golovenko, D., Demo, G., Dutcher, R.C., Tanaka Hall, T.M., Zamore, P.D. et al. (2022) Structural basis of MicroRNA biogenesis by Dicer-1 and its partner protein Loqs-PB. *Mol. Cell.* **82**, 4049–4063.E6, https://doi.org/10.1016/j.molcel.2022.09.002
125 Nguyen, T.A., Jo, M.H., Choi, Y.G., Park, J., Kwon, S.C., Hohng, S. et al. (2015) Functional anatomy of the human microprocessor. *Cell* **161**, 1374–1387, https://doi.org/10.1016/j.cell.2015.05.010
126 Rice, G.M., Shivashankar, V., Ma, E.J., Baryza, J.L. and Nutiu, R. (2020) Functional Atlas of primary miRNA maturation by the microprocessor. *Mol. Cell.* **80**, 892e4–902e4, https://doi.org/10.1016/j.molcel.2020.10.028
127 Jin, W., Wang, J., Liu, C.P., Wang, H.W. and Xu, R.M. (2020) Structural basis for pri-miRNA recognition by Drosha. *Mol. Cell.* **78**, 423e5–433e5, https://doi.org/10.1016/j.molcel.2020.02.024
128 Partin, A.C., Zhang, K., Jeong, B.C., Herrell, E., Li, S., Chiu, W. et al. (2020) Cryo-EM structures of human Drosha and DGCR8 in complex with primary microRNA. *Mol. Cell.* **78**, 411e4–422e4, https://doi.org/10.1016/j.molcel.2020.02.016
129 Saldi, T., Fong, N. and Bentley, D.L. (2018) Transcription elongation rate affects nascent histone pre-mRNA folding and 3′ end processing. *Genes Dev.* **32**, 297–308, https://doi.org/10.1101/gad.310896.117
130 Muniz, L., Nicolas, E. and Trouche, D. (2021) RNA polymerase II speed: a key player in controlling and adapting transcriptome composition. *EMBO J.* **40**, e105740, https://doi.org/10.15252/embj.2020105740
131 Wissink, E.M., Vihervaara, A., Tippens, N.D. and Lis, J.T. (2019) Nascent RNA analyses: tracking transcription and its regulation. *Nat. Rev. Genet.* **20**, 705–723, https://doi.org/10.1038/s41576-019-0159-6
132 Eick, D. and Geyer, M. (2013) The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem. Rev.* **113**, 8456–8490, https://doi.org/10.1021/cr400071f
133 Szlachta, K., Thys, R.G., Atkin, N.D., Pierce, L.C.T., Bekiranov, S. and Wang, Y.H. (2018) Alternative DNA secondary structure formation affects RNA polymerase II promoter-proximal pausing in human. *Genome Biol.* **19**, 89, https://doi.org/10.1186/s13059-018-1463-8
134 Zumer, K., Maier, K.C., Farnung, L., Jaeger, M.G., Rus, P., Winter, G. et al. (2021) Two distinct mechanisms of RNA polymerase II elongation stimulation in vivo. *Mol. Cell.* **81**, 3096e8–3109e8, https://doi.org/10.1016/j.molcel.2021.05.028
135 Boccaletto, P., Stefaniak, F., Ray, A., Cappannini, A., Mukherjee, S., Purta, E. et al. (2022) MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res.* **50**, D231–D235, https://doi.org/10.1093/nar/gkab1083
136 Shi, H., Wei, J. and He, C. (2019) Where, when, and how: context-dependent functions of RNA methylation writers, readers, and erasers. *Mol. Cell.* **74**, 640–650, https://doi.org/10.1016/j.molcel.2019.04.025
137 Borchardt, E.K., Martinez, N.M. and Gilbert, W.V. (2020) Regulation and function of RNA Pseudouridylation in Human Cells. *Annu. Rev. Genet.* **54**, 309–336, https://doi.org/10.1146/annurev-genet-112618-043830
138 Roost, C., Lynch, S.R., Batista, P.J., Qu, K., Chang, H.Y. and Kool, E.T. (2015) Structure and thermodynamics of N6-methyladenosine in RNA: a spring-loaded base modification. *J. Am. Chem. Soc.* **137**, 2107–2115, https://doi.org/10.1021/ja513080v
139 Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M. and Pan, T. (2015) N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560–564, https://doi.org/10.1038/nature14234
140 Deb, I., Popenda, L., Sarzynska, J., Malgowska, M., Lahiri, A., Gdaniec, Z. et al. (2019) Computational and NMR studies of RNA duplexes with an internal pseudouridine-adenosine base pair. *Sci. Rep.* **9**, 16278, https://doi.org/10.1038/s41598-019-52637-0
141 Zhao, Y., Rai, J., Yu, H. and Li, H. (2022) CryoEM structures of pseudouridine-free ribosome suggest impacts of chemical modifications on ribosome conformations. *Structure* **30**, 983–92e5
142 Martinez, N.M., Su, A., Burns, M.C., Nussbacher, J.K., Schaening, C., Sathe, S. et al. (2022) Pseudouridine synthases modify human pre-mRNA co-transcriptionally and affect pre-mRNA processing. *Mol. Cell.* **82**, 645e9–659e9, https://doi.org/10.1016/j.molcel.2021.12.023
143 Ke, S., Pandya-Jones, A., Saito, Y., Fak, J.J., Vagbo, C.B., Geula, S. et al. (2017) m(6)A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. *Genes Dev.* **31**, 990–1006, https://doi.org/10.1101/gad.301036.117
144 Wei, G., Almeida, M., Pintacuda, G., Coker, H., Bowness, J.S., Ule, J. et al. (2021) Acute depletion of METTL3 implicates N (6)-methyladenosine in alternative intron/exon inclusion in the nascent transcriptome. *Genome Res.* **31**, 1395–1408, https://doi.org/10.1101/gr.271635.120
145 De Bortoli, F., Espinosa, S. and Zhao, R. (2021) DEAH-box RNA helicases in pre-mRNA splicing. *Trends Biochem. Sci.* **46**, 225–238, https://doi.org/10.1016/j.tibs.2020.10.006
146 Bohnsack, K.E., Ficner, R., Bohnsack, M.T. and Jonas, S. (2021) Regulation of DEAH-box RNA helicases by G-patch proteins. *Biol. Chem.* **402**, 561–579, https://doi.org/10.1515/hsz-2020-0338
147 Song, H. and Ji, X. (2019) The mechanism of RNA duplex recognition and unwinding by DEAD-box helicase DDX3X. *Nat. Commun.* **10**, 3085, https://doi.org/10.1038/s41467-019-11083-2
148 Umate, P., Tuteja, N. and Tuteja, R. (2011) Genome-wide comprehensive analysis of human helicases. *Commun Integr Biol.* **4**, 118–137, https://doi.org/10.4161/cib.13844
149 England, W.E., Wang, J., Chen, S., Baldi, P., Flynn, R.A. and Spitale, R.C. (2022) An atlas of posttranslational modifications on RNA binding proteins. *Nucleic Acids Res.* **50**, 4329–39, https://doi.org/10.1093/nar/gkac243
150 Cho, S. and Dreyfuss, G. (2010) A degron created by SMN2 exon 7 skipping is a principal contributor to spinal muscular atrophy severity. *Genes Dev.* **24**, 438–442, https://doi.org/10.1101/gad.1884910

151 Kino, Y., Washizu, C., Kurosawa, M., Oma, Y., Hattori, N., Ishiura, S. et al. (2015) Nuclear localization of MBNL1: splicing-mediated autoregulation and repression of repeat-derived aberrant proteins. *Hum. Mol. Genet.* **24**, 740–756, https://doi.org/10.1093/hmg/ddu492
152 Tran, H., Gourrier, N., Lemercier-Neuillet, C., Dhaenens, C.M., Vautrin, A., Fernandez-Gomez, F.J. et al. (2011) Analysis of exonic regions involved in nuclear localization, splicing activity, and dimerization of Muscleblind-like-1 isoforms. *J. Biol. Chem.* **286**, 16435–16446, https://doi.org/10.1074/jbc.M110.194928
153 Wang, E.T., Cody, N.A., Jog, S., Biancolella, M., Wang, T.T., Treacy, D.J. et al. (2012) Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **150**, 710–724, https://doi.org/10.1016/j.cell.2012.06.041
154 Malik, I., Kelley, C.P., Wang, E.T. and Todd, P.K. (2021) Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat. Rev. Mol. Cell Biol.* **22**, 589–607, https://doi.org/10.1038/s41580-021-00382-6
155 Guan, F., Caratozzolo, R.M., Goraczniak, R., Ho, E.S. and Gunderson, S.I. (2007) A bipartite U1 site represses U1A expression by synergizing with PIE to inhibit nuclear polyadenylation. *RNA* **13**, 2129–2140, https://doi.org/10.1261/rna.756707
156 Barash, Y., Calarco, J.A., Gao, W., Pan, Q., Wang, X., Shai, O. et al. (2010) Deciphering the splicing code. *Nature* **465**, 53–59, https://doi.org/10.1038/nature09000
157 Ciolli Mattioli, C., Rom, A., Franke, V., Imami, K., Arrey, G., Terne, M. et al. (2019) Alternative 3′ UTRs direct localization of functionally diverse protein isoforms in neuronal compartments. *Nucleic Acids Res.* **47**, 2560–2573, https://doi.org/10.1093/nar/gky1270
158 Corley, M., Solem, A., Phillips, G., Lackey, L., Ziehr, B., Vincent, H.A. et al. (2017) An RNA structure-mediated, posttranscriptional model of human alpha-1-antitrypsin expression. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E10244–E10253, https://doi.org/10.1073/pnas.1706539114
159 Scharner, J. and Aznarez, I. (2021) Clinical applications of single-stranded oligonucleotides: current landscape of approved and in-development therapeutics. *Mol. Ther.* **29**, 540–554, https://doi.org/10.1016/j.ymthe.2020.12.022
160 Kole, R., Krainer, A.R. and Altman, S. (2012) RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov.* **11**, 125–140, https://doi.org/10.1038/nrd3625
161 Olson, S.W., Turner, A.W., Arney, J.W., Saleem, I., Weidmann, C.A., Margolis, D.M. et al. (2022) Discovery of a large-scale, cell-state-responsive allosteric switch in the 7SK RNA using DANCE-MaP. *Mol. Cell.* **82**, 1708e10–1723e10, https://doi.org/10.1016/j.molcel.2022.02.009
162 Lan, T.C.T., Allan, M.F., Malsick, L.E., Woo, J.Z., Zhu, C., Zhang, F. et al. (2022) Secondary structural ensembles of the SARS-CoV-2 RNA genome in infected cells. *Nat. Commun.* **13**, 1128, https://doi.org/10.1038/s41467-022-28603-2
163 Quemener, A.M. and Galibert, M.D. (2021) Antisense oligonucleotide: A promising therapeutic option to beat COVID-19. *Wiley Interdiscip. Rev. RNA* e1703
164 Green, A.A., Silver, P.A., Collins, J.J. and Yin, P. (2014) Toehold switches: de-novo-designed regulators of gene expression. *Cell* **159**, 925–939, https://doi.org/10.1016/j.cell.2014.10.002
165 Pardee, K., Green, A.A., Takahashi, M.K., Braff, D., Lambert, G., Lee, J.W. et al. (2016) Rapid, low-cost detection of Zika virus using programmable biomolecular components. *Cell* **165**, 1255–1266, https://doi.org/10.1016/j.cell.2016.04.059
166 Park, S. and Lee, J.W. (2021) Detection of coronaviruses using RNA toehold switch sensors. *Int. J. Mol. Sci.* **22**, 1772, https://doi.org/10.3390/ijms22041772
167 Hong, F., Ma, D., Wu, K., Mina, L.A., Luiten, R.C., Liu, Y. et al. (2020) Precise and programmable detection of mutations using ultraspecific riboregulators. *Cell* **183**, 835–836, https://doi.org/10.1016/j.cell.2020.10.020
168 Zadeh, J.N., Steenberg, C.D., Bois, J.S., Wolfe, B.R., Pierce, M.B., Khan, A.R. et al. (2011) NUPACK: analysis and design of nucleic acid systems. *J. Comput. Chem.* **32**, 170–173, https://doi.org/10.1002/jcc.21596
169 Glasscock, C.J., Biggs, B.W., Lazar, J.T., Arnold, J.H., Burdette, L.A., Valdes, A. et al. (2021) Dynamic control of gene expression with riboregulated switchable feedback promoters. *ACS Synth Biol.* **10**, 1199–1213, https://doi.org/10.1021/acssynbio.1c00015
170 Warner, K.D., Hajdin, C.E. and Weeks, K.M. (2018) Principles for targeting RNA with drug-like small molecules. *Nat. Rev. Drug Discov.* **17**, 547–558, https://doi.org/10.1038/nrd.2018.93
171 Falese, J.P., Donlic, A. and Hargrove, A.E. (2021) Targeting RNA with small molecules: from fundamental principles towards the clinic. *Chem. Soc. Rev.* **50**, 2224–2243, https://doi.org/10.1039/D0CS01261K
172 Meyer, S.M., Williams, C.C., Akahori, Y., Tanaka, T., Aikawa, H., Tong, Y. et al. (2020) Small molecule recognition of disease-relevant RNA structures. *Chem. Soc. Rev.* **49**, 7167–7199, https://doi.org/10.1039/D0CS00560F
173 Serganov, A. and Nudler, E. (2013) A decade of riboswitches. *Cell* **152**, 17–24, https://doi.org/10.1016/j.cell.2012.12.024
174 Aguilar, R., Spencer, K.B., Kesner, B., Rizvi, N.F., Badmalia, M.D., Mrozowich, T. et al. (2022) Targeting Xist with compounds that disrupt RNA structure and X inactivation. *Nature* **604**, 160–166, https://doi.org/10.1038/s41586-022-04537-z
175 Donlic, A., Morgan, B.S., Xu, J.L., Liu, A., Roble, Jr., C. and Hargrove, A.E. (2018) Discovery of small molecule ligands for MALAT1 by tuning an RNA-binding scaffold. *Angew. Chem. Int. Ed. Engl.* **57**, 13242–13247, https://doi.org/10.1002/anie.201808823
176 Singh, R.N., Ottesen, E.W. and Singh, N.N. (2020) The first orally deliverable small molecule for the treatment of spinal muscular atrophy. *Neurosci. Insights* **15**, 2633105520973985, https://doi.org/10.1177/2633105520973985
177 Campagne, S., Boigner, S., Rudisser, S., Moursy, A., Gillioz, L., Knorlein, A. et al. (2019) Structural basis of a small molecule targeting RNA for a specific splicing correction. *Nat. Chem. Biol.* **15**, 1191–1198, https://doi.org/10.1038/s41589-019-0384-5
178 Palacino, J., Swalley, S.E., Song, C., Cheung, A.K., Shu, L., Zhang, X. et al. (2015) SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nat. Chem. Biol.* **11**, 511–517, https://doi.org/10.1038/nchembio.1837
179 Wells, S.E., Hughes, J.M., Igel, A.H. and Ares, Jr, M. (2000) Use of dimethyl sulfate to probe RNA structure in vivo. *Methods Enzymol.* **318**, 479–493, https://doi.org/10.1016/S0076-6879(00)18071-1
180 Tijerina, P., Mohr, S. and Russell, R. (2007) DMS footprinting of structured RNAs and RNA-protein complexes. *Nat. Protoc.* **2**, 2608–2623, https://doi.org/10.1038/nprot.2007.380

181 Guo, L.T., Adams, R.L., Wan, H., Huston, N.C., Potapova, O., Olson, S. et al. (2020) Sequencing and structure probing of long RNAs using MarathonRT: a next-generation reverse transcriptase. *J. Mol. Biol.* **432**, 3338–3352, https://doi.org/10.1016/j.jmb.2020.03.022
182 Poulsen, L.D., Kielpinski, L.J., Salama, S.R., Krogh, A. and Vinther, J. (2015) SHAPE selection (SHAPES) enrich for RNA structure signal in SHAPE sequencing-based probing data. *RNA* **21**, 1042–1052, https://doi.org/10.1261/rna.047068.114
183 Corley, M., Flynn, R.A., Blue, S.M., Yee, B.A., Chang, H.Y. and Yeo, G.W. (2021) fSHAPE, fSHAPE-eCLIP, and SHAPE-eCLIP probe transcript regions that interact with specific proteins. *STAR Protoc.* **2**, 100762, https://doi.org/10.1016/j.xpro.2021.100762
184 Corley, M., Flynn, R.A., Lee, B., Blue, S.M., Chang, H.Y. and Yeo, G.W. (2020) Footprinting SHAPE-eCLIP reveals transcriptome-wide hydrogen bonds at RNA-protein interfaces. *Mol. Cell.* **80**, 903e8–914e8, https://doi.org/10.1016/j.molcel.2020.11.014
185 Lu, Z., Zhang, Q.C., Lee, B., Flynn, R.A., Smith, M.A., Robinson, J.T. et al. (2016) RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* **165**, 1267–1279, https://doi.org/10.1016/j.cell.2016.04.028
186 Helwak, A. and Tollervey, D. (2014) Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH). *Nat. Protoc.* **9**, 711–728, https://doi.org/10.1038/nprot.2014.043
187 Aw, J.G., Shen, Y., Wilm, A., Sun, M., Lim, X.N., Boon, K.L. et al. (2016) In vivo mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. *Mol. Cell.* **62**, 603–617, https://doi.org/10.1016/j.molcel.2016.04.028
188 Sharma, E., Sterne-Weiler, T., O'Hanlon, D. and Blencowe, B.J. (2016) Global mapping of human RNA-RNA interactions. *Mol. Cell.* **62**, 618–626, https://doi.org/10.1016/j.molcel.2016.04.030
189 Cao, C., Cai, Z., Ye, R., Su, R., Hu, N., Zhao, H. et al. (2021) Global in situ profiling of RNA-RNA spatial interactions with RIC-seq. *Nat. Protoc.* **16**, 2916–2946, https://doi.org/10.1038/s41596-021-00524-2
190 Yoshida, H., Matsui, T., Yamamoto, A., Okada, T. and Mori, K. (2001) XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell* **107**, 881–891, https://doi.org/10.1016/S0092-8674(01)00611-0
191 Warf, M.B. and Berglund, J.A. (2007) MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA* **13**, 2238–2251, https://doi.org/10.1261/rna.610607
192 Buratti, E., Dhir, A., Lewandowska, M.A. and Baralle, F.E. (2007) RNA structure is a key regulatory element in pathological ATM and CFTR pseudoexon inclusion events. *Nucleic Acids Res.* **35**, 4369–4383, https://doi.org/10.1093/nar/gkm447
193 Li, G., Shen, J., Cao, J., Zhou, G., Lei, T., Sun, Y. et al. (2018) Alternative splicing of human telomerase reverse transcriptase in gliomas and its modulation mediated by CX-5461. *J. Exp. Clin. Cancer Res.* **37**, 78, https://doi.org/10.1186/s13046-018-0749-8
194 Wong, M.S., Shay, J.W. and Wright, W.E. (2014) Regulation of human telomerase splicing by RNA:RNA pairing. *Nat. Commun.* **5**, 3306, https://doi.org/10.1038/ncomms4306
195 Marcel, V., Tran, P.L., Sagne, C., Martel-Planche, G., Vaslin, L., Teulade-Fichou, M.P. et al. (2011) G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis* **32**, 271–278, https://doi.org/10.1093/carcin/bgq253
196 Blice-Baum, A.C. and Mihailescu, M.R. (2014) Biophysical characterization of G-quadruplex forming FMR1 mRNA and of its interactions with different fragile X mental retardation protein isoforms. *RNA* **20**, 103–114, https://doi.org/10.1261/rna.041442.113
197 Huang, H., Zhang, J., Harvey, S.E., Hu, X. and Cheng, C. (2017) RNA G-quadruplex secondary structure promotes alternative splicing via the RNA-binding protein hnRNPF. *Genes Dev.* **31**, 2296–2309, https://doi.org/10.1101/gad.305862.117
198 Lovci, M.T., Ghanem, D., Marr, H., Arnold, J., Gee, S., Parra, M. et al. (2013) Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. *Nat. Struct. Mol. Biol.* **20**, 1434–1442, https://doi.org/10.1038/nsmb.2699
199 Taube, J.R., Sperle, K., Banser, L., Seeman, P., Cavan, B.C., Garbern, J.Y. et al. (2014) PMD patient mutations reveal a long-distance intronic interaction that regulates PLP1/DM20 alternative splicing. *Hum. Mol. Genet.* **23**, 5464–5478, https://doi.org/10.1093/hmg/ddu271
200 Kalinina, M., Skvortsov, D., Kalmykova, S., Ivanov, T., Dontsova, O. and Pervouchine, D.D. (2021) Multiple competing RNA structures dynamically control alternative splicing in the human ATE1 gene. *Nucleic Acids Res.* **49**, 479–490, https://doi.org/10.1093/nar/gkaa1208
201 Higashide, S., Morikawa, K., Okumura, M., Kondo, S., Ogata, M., Murakami, T. et al. (2004) Identification of regulatory cis-acting elements for alternative splicing of presenilin 2 exon 5 under hypoxic stress conditions. *J. Neurochem.* **91**, 1191–1198, https://doi.org/10.1111/j.1471-4159.2004.02798.x
202 Kralovicova, J., Patel, A., Searle, M. and Vorechovsky, I. (2015) The role of short RNA loops in recognition of a single-hairpin exon derived from a mammalian-wide interspersed repeat. *RNA Biol.* **12**, 54–69, https://doi.org/10.1080/15476286.2015.1017207
203 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S. et al. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649, https://doi.org/10.1093/bioinformatics/bts199
204 Wang, J., Schultz, P.G. and Johnson, K.A. (2018) Mechanistic studies of a small-molecule modulator of SMN2 splicing. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E4604–E4612
205 Busan, S. and Weeks, K.M. (2018) Accurate detection of chemical modifications in RNA by mutational profiling (MaP) with ShapeMapper 2. *RNA* **24**, 143–148, https://doi.org/10.1261/rna.061945.117
206 Mathews, D.H. (2004) Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. *RNA* **10**, 1178–1190, https://doi.org/10.1261/rna.7650904
207 Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. et al. (2011) Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26, https://doi.org/10.1038/nbt.1754