

E2,3ELL
HO H222



Molecular Biology of RNA

David Elliott

Institute of Human Genetics, Newcastle University

Michael Ladomery

Faculty of Health and Life Sciences, University of the West of England, Bristol

OXFORD
UNIVERSITY PRESS



CONTENTS

<i>Preface</i>	v		
<i>Acknowledgements</i>	x		
1 Introduction to <i>Molecular Biology of RNA</i>			
1.1 Aims of this book	2		
2 RNA can form versatile structures			
2.1 DNA and RNA are composed of slightly different building blocks	4		
2.2 Nucleotides are joined together through a phosphodiester backbone to give nucleotide chains	8		
2.3 RNA secondary structure: hydrogen bonding between bases holds nucleotide chains together in double helices	9		
2.4 Nucleic acids have primary, secondary, and, in the case of RNA, tertiary structure	12		
2.5 Five common secondary structure motifs are found within RNA molecules	14		
2.6 Secondary structure motifs form through base pairing in RNA molecules	16		
2.7 The formation of RNA duplexes is stimulated by positively charged molecules and particularly metal ions	19		
2.8 RNAs form tertiary structures	20		
2.9 Complex folded RNAs which bind to target molecules can be selected	24		
2.10 Riboswitches are shape-changing RNAs which can flip gene expression patterns on binding specific target molecules	25		
2.11 RNA helices can connect different molecules of RNA together	30		
2.12 RNA is a versatile molecule	32		
3 Catalytic RNAs			
3.1 RNA is inherently chemically unstable because of its 2' -OH group	35	3.4 Ribozymes are widespread in nature and fall into large and small groups	40
3.2 The protein enzyme RNase A uses acid-base catalysis to carry out RNA strand cleavage	37	3.5 Small ribonucleolytic ribozymes catalyse their own cleavage	43
3.3 Three properties of RNA enable the catalytic function of ribozymes	39	3.6 The hammerhead ribozyme	43
		3.7 The HDV ribozyme	46
		3.8 RNA-cutting ribozymes are used to control gene expression in both bacteria and eukaryotes	46
		3.9 The large ribozymes	47
		3.10 Group I introns are spliced through a two-step mechanism which uses metal ions in their active sites	47
		3.11 Group II introns are also spliced through a two-step mechanism	51
		3.12 RNase P is an essential ribozyme which processes the 5' end of tRNA	53
		3.13 Catalysis in the ribosome is RNA based	53
		3.14 Are ribozymes true catalysts?	56
		3.15 The RNA World hypothesis: a time when RNA was used as a genetic material	57
		3.16 Experiments have been carried out to model the early steps that might have occurred during the evolution of life	58
		4 The RNA-binding proteins	
		4.1 The RNA recognition motif (RRM)	66
		4.2 The K-homology (KH) domain	69
		4.3 The cold-shock domain	70
		4.4 Double-stranded RNA-binding proteins	73
		4.5 The zinc-finger domain	75
		4.6 Other RNA-binding domains	77
		4.7 Investigating protein-RNA interactions	79
		5 Co-transcriptional pre-mRNA processing	
		5.1 Transcription and the RNA polymerases	85
		5.2 Formation of the ends of an mRNA	88
		5.3 The C-terminal domain (CTD) of RNA polymerase II	91
		5.4 The link between splicing and transcription	92

5.5 The spatial organization of pre-mRNA processing	96	the splicing machinery, thereby directing exons to be spliced	145
5.6 Histone mRNA 3' end formation	100	7.8 Negative splicing complexes exclude assembly of the spliceosome	148
6 Pre-mRNA splicing by the spliceosome		7.9 How the splicing code is being deciphered	150
6.1 RNA splicing was discovered in a virus	104	7.10 The splicing code can be predicted using bioinformatics	151
6.2 Spliceosomal introns play a critical role in efficient eukaryotic gene expression	106	7.11 Experimental identification of splicing silencers	151
6.3 Introns enhance eukaryotic gene expression at several levels	108	7.12 Most exons contain multiple binding sites for antagonistic splicing factors	153
6.4 Pre-mRNAs are punctuated by splice sites at intron-exon junctions	109	7.13 Four features affect how pre-mRNAs are coded by the spliceosome	154
6.5 Splice sites are complementary to a group of small nuclear RNAs	111	8 Regulated RNA processing—alternative splicing	
6.6 snRNAs are associated with proteins to give snRNPs	112	8.1 Alternative splicing produces different mRNA isoforms from the same gene	159
6.7 Splicing follows a two-step reaction pathway	114	8.2 The different mRNA isoforms made by alternative splicing can be detected by RT-PCR	160
6.8 Splicing happens in a series of spliceosomal protein complexes	115	8.3 Alternatively spliced mRNA isoforms can be displayed on genome browsers	161
6.9 The spliceosome cycle	117	8.4 Alternative splicing increases the coding capacity of the genome by challenging the 'one gene—one protein' rule	162
6.10 Spliceosome assembly and disassembly are cyclical	121	8.5 Alternative splicing can regulate both protein-coding information and gene expression	164
6.11 A minor class of eukaryotic spliceosomal introns have different splice sites	123	8.6 Changes in the splicing code can regulate alternative splicing patterns	167
6.12 Major and minor spliceosomes coexist in most eukaryotes	124	8.7 Signal transduction pathways can regulate alternative splicing by changing the function and location of splicing factors	172
6.13 <i>Trans</i> -splicing is common in trypanosome parasites and in the nematode <i>C. elegans</i> , where it enables efficient translation	124	8.8 Protein phosphorylation	172
6.14 Pre-mRNA splicing is thought to have evolved from parasitic DNA elements	128	8.9 Stimulation of cells with growth factors switches the splicing of the cell surface molecule CD44	173
6.15 Introns evolved early in eukaryotes	129	8.10 The splicing repressor hnRNP A1 relocalizes to the cytoplasm in response to cellular stress	174
6.16 Introns and exons can both appear and disappear in evolution	130	8.11 Splicing decisions can be regulated by dephosphorylation of splicing factors	177
7 How pre-mRNAs are decoded by the splicing machinery		8.12 Transcription elongation speeds can regulate alternative splicing choices	178
7.1 Intron and exon definition	137	8.13 Rates of elongation of RNA polymerase II can be regulated to affect alternative splicing	180
7.2 A splicing code helps exon recognition and so controls splicing	140	8.14 Transcription can also modulate splicing pathways via the recruitment of cofactors	181
7.3 Discovery of the splicing code	140		
7.4 The splicing code comprises binding sites for nuclear RNA-binding proteins embedded in transcribed sequences	141		
7.5 The <i>cis</i> -splicing code is read by nuclear RNA-binding proteins	142		
7.6 Splicing enhancers bind proteins to activate the use of splice sites	145		
7.7 Proteins bound to splicing enhancers act as patches of glue on the RNA for attaching			

145	8.15 Alternative splicing in action: alternative splicing pathways can control complex developmental pathways in metazoans	183	10 Nucleocytoplasmic traffic of messenger RNA
148			10.1 Common themes in RNA nuclear export pathways
150	9 Pre-mRNA splicing defects in development and disease		231
151	9.1 Splicing mutations are very frequent causes of human genetic disease	193	10.2 The nuclear pore complex (NPC)
151	9.2 Mutations in splicing control sequences frequently cause exon skipping in humans	194	232
153	9.3 Molecular diagnosis of splicing mutations	195	10.3 Nuclear export of mRNA
154	9.4 Mutation of an exonic splicing enhancer in a DNA damage control gene leads to breast cancer	197	233
g	9.5 Genetic mutations create a new splice site in a premature ageing disease	200	10.4 The cell biology of nuclear export of mRNA: mRNA transcripts reach the nuclear pore by random nuclear diffusion
159	9.6 Mutations which affect splicing can deregulate the ratio of alternatively spliced mRNAs	201	234
160	9.7 Mutations affecting splicing signals can be particularly severe since they change the structure of mRNAs	203	10.5 As they are synthesized, mRNAs are 'dressed for export' into mRNPs by the addition of nuclear export adaptors
161	9.8 Manipulating pre-mRNA splicing offers a route to treating muscular dystrophy	205	236
162	9.9 Diseases caused by mutations in the <i>trans</i> -acting machinery which recognizes and splices together exons in the nucleus	207	10.6 The molecular mechanism of RNA export
164	9.10 The genes which encode important spliceosomal proteins are mutated in patients with retinitis pigmentosa (RP)	208	239
167	9.11 A protein important for snRNP assembly is affected by mutations causing spinal muscular atrophy (SMA)	211	10.7 Export adaptors
172	9.12 There is scientific controversy about why exon 7 of the <i>SMN2</i> gene is inefficiently spliced	212	239
172	9.13 Molecular therapy for SMA is targeted at correcting the splicing of <i>SMN2</i> exon 7	214	10.8 Addition of export adaptors is coupled to splicing in metazoans and transcription in yeast
173	9.14 Diseases caused by mis-expression of levels of splicing factor	216	244
174	9.15 The splicing regulator ASF/SF2 is misregulated in some cancers	216	10.9 Nuclear export receptors move export-competent mRNPs through the nuclear pore
177	9.16 Myotonic dystrophy is caused by the expression of pathogenic RNAs which affect pre-mRNA splicing by changing the nuclear concentration of nuclear pre-mRNA splicing proteins	219	246
178	9.17 Splicing proteins can play an important role in autoimmunity	223	10.10 The logic of mRNA export
80	9.18 Splicing as a route to therapy for infectious diseases like AIDS	225	247
81			10.11 Hijacking of the mRNA export machinery: the constitutive transport element sequence directs the nuclear export of unspliced transcripts from the MPMV virus
			247
			10.12 Movement of mRNP through the nuclear pore
			248
			11 Nucleocytoplasmic traffic of non-coding RNA
			11.1 Compartment-specific transport complexes
			252
			11.2 Nuclear transport of rRNA, tRNA, snRNAs, and microRNAs is dependent on the RAN GTPase protein
			253
			11.3 Different forms of RAN are found in the nucleus and cytoplasm
			253
			11.4 Non-coding RNA nuclear export complexes contain adaptors and receptors
			254
			11.5 Nuclear export of ncRNA is dependent on nuclear export adaptor and receptor proteins
			255
			11.6 Karyopherins are an important group of nuclear export receptors which respond to positional information provided by RAN
			256
			11.7 CRM1 is the nuclear export receptor (karyopherin) for rRNAs and snRNAs
			258
			11.8 Different karyopherins act as export receptors for tRNA and microRNAs
			259
			11.9 After releasing their loads in the cytoplasm nuclear RNA export components are moved back into the nucleus
			259

11.10	Nuclear transport of snRNAs	261	15.5	Alu elements are the main targets of A → I editing in humans	327
11.11	Mature snRNPs are re-imported into the nucleus using the nuclear protein import machinery	263	15.6	Selective A → I RNA editing by ADAR enzymes modifies mRNAs that contain short regions of dsRNA	329
11.12	Mature U6 snRNP is made exclusively in the nucleus	265	15.7	The four known biological functions of A → I mRNA editing in the cell	329
11.13	Retroviruses have hijacked the RNA export machinery to assist in the export of partially processed mRNAs	266	15.8	ADAR proteins are essential for normal nervous system development, but also play roles elsewhere in the body	334
12	Messenger RNA localization		15.9	A → I editing plays an important role in the function of tRNAs	335
12.1	The need for mRNA localization	270	15.10	C → U RNA editing takes place through base deamination (removal of an amino group from cytidine)	336
12.2	The machinery of mRNA localization	272	15.11	C → U RNA editing makes two different forms of the APOB mRNA in different tissues, and was the first RNA editing reaction to be discovered in animals	336
12.3	Classical examples of mRNA localization in development	274	15.12	An RNA editing complex containing the cytosine deaminase ApoBec1 binds and edits APOB mRNA	337
12.4	Localization of mRNA in differentiated somatic cells	277	15.13	ApoBec proteins play an important role in innate immunity to retroviruses like HIV and in generating an antibody response	339
12.5	Localization of mRNA in plants	281	15.14	Trypanosome mitochondrial RNA is edited by base insertions and deletions to create ORFs from frameshifted transcripts	342
13	Translation of messenger RNA		15.15	RNA editing was discovered in trypanosomes by sequencing cDNAs encoded by mitochondrial genes	343
13.1	What is translation?	285	15.16	Short RNAs called guide RNAs target trypanosome mitochondrial RNA editing	345
13.2	The structure of the ribosome	285	15.17	Guide RNAs are used as a template for RNA editing through uridine insertions and deletions	346
13.3	Deciphering the genetic code	287	15.18	Trypanosome mitochondrial RNA editing requires nuclear-encoded proteins which might be useful therapeutic targets	346
13.4	Three key steps in translation	290	16	The biogenesis of non-coding RNAs	
13.5	Regulation of mRNA translation	292	16.1	The snoRNAs and scaRNAs: multiple roles in RNA biogenesis	352
13.6	The masked messages	299	16.2	Structure and function of the nucleolus	358
13.7	Manipulating translation	302	16.3	Processing of tRNA and of mitochondrial transcripts	361
14	Stability and degradation of mRNA				
14.1	Messenger RNAs have a half-life	307			
14.2	Sites and mechanisms of mRNA degradation	308			
14.3	The process of mRNA degradation	309			
14.4	Extracellular stimuli influence the stability of mRNA	314			
14.5	Nonsense-mediated mRNA decay	316			
14.6	Degradation of mRNA in bacteria and plants	319			
15	RNA editing				
15.1	Why edit RNA?	322			
15.2	A → I editing takes place by modification of adenosine through removal of an amino group	323			
15.3	A → I editing affects RNA hydrogen bonding between bases since inosine forms stable base pairs with cytosine	325			
15.4	A → I editing was discovered because it destabilized dsRNAs	326			

327	17 The 'mega RNAs': long non-coding RNAs with important roles in epigenetic regulation of gene expression		
329	17.1 Introduction to epigenetic regulatory ncRNAs	366	
329	17.2 Transcriptionally active and inactive DNA is created by tagging chromatin with simple chemical groups	368	
334	17.3 Long ncRNAs help epigenetically programme important developmental control genes	374	
335	17.4 A difference in size of the sex chromosomes means there is a requirement for dosage compensation	377	
336	17.5 Dosage compensation in female mammals uses non-coding RNAs to inactivate one female X chromosome	377	
336	17.6 Dosage compensation in fruit flies uses a dosage compensation complex including a long ncRNA to up-regulate expression from a single male X chromosome	381	
337	17.7 Similarities and differences between mechanisms of dosage compensation in fruit flies and mammals	382	
/ and			
339			
342			
343			
345			
346			
346			
352			
358			
361			
	17.8 Genetic imprinting controls gene expression depending on the parent of origin of the gene or chromosome		384
	17.9 Parentally imprinted gene clusters often include long ncRNAs		385
	17.10 The ncRNA <i>AIR</i> epigenetically represses <i>IGF2R</i> gene expression by directing epigenetic chromatin modification		385
	17.11 Transcription of <i>H19</i> ncRNA acts as a decoy for transcription of the <i>IGF2</i> gene		387
	18 The short non-coding RNAs and gene silencing		
	18.1 Key concepts and common pathways		391
	18.2 Discovery and mechanism of RNA interference		394
	18.3 The uses of RNA interference		398
	18.4 Discovery, biogenesis, and developmental roles of microRNAs		403
	18.5 Transcriptional silencing by non-coding RNAs in the centromere		408
	18.6 RNA-induced transcriptional silencing of transposons		411
	<i>Glossary</i>		421