

The relationship between miRNA biogenesis and RNA splicing

Aimee Parsons

**Thesis submitted to the University of Nottingham for the
degree of Doctor of Philosophy**

September 2017

Declaration

Except where acknowledged in the text, I declare that this dissertation is my own work and it is based on the research work that was undertaken by me in Gene Regulation & RNA Biology group, School of Pharmacy, Faculty of Science, University of Nottingham, UK.

Abstract

MiRNAs (miRNAs) are important for the proper regulation of translation, with aberrant miRNA expression contributing to diseases such as cancer. MiRNAs are expressed from longer primary-miRNA transcripts and found in a variety of genomic locations, including introns and exons of coding or long noncoding RNAs. MiRNA biogenesis begins with excision of the precursor miRNA hairpin by the Microprocessor complex which consists of two proteins: Drosha and DGCR8. Both Microprocessor cleavage and splicing occur co-transcriptionally.

Splicing and Microprocessing coexist when a miRNA hairpin is located within an intron without detriment to mature mRNA or miRNA production, but little is known about how these two processes interact when a miRNA hairpin is located within an exon. Similarly, little is known about how a miRNA is processed from a long non-coding (lnc)RNA, which is an RNA transcript longer than 200 nucleotides that does not code for a protein. Intronic miRNAs are processed without detriment to the splicing of the host transcript and production of functional mRNA and protein; however, exonic miRNA processing would be expected to lead to cleavage of the exon, interfering with production of mature mRNA. To understand how the genomic location of a miRNA hairpin affects the splicing of the host gene and how splicing affects miRNA production from different locations, two different approaches were used.

Firstly, liver-specific miR-122 was inserted at a series of locations within a β -globin plasmid and transfected into HeLa cells, which do not express miR-122. qPCR and northern blotting show that mature miR-122 expression is increased when located within either an intron or exon compared to its endogenous context within a long noncoding RNA. Importantly, the levels of spliced β -globin was shown to decrease when miR-122 was expressed from an exon. This suggests that there is some competition between miRNA biogenesis and splicing.

Secondly, the effect of the splice inhibitor pla B was used to understand the effect of splicing on endogenous pri-miR-122. Splice inhibition led to a decrease in the levels of spliced pri-miR-122, but interestingly also led to a

decrease in the levels of unspliced transcript indicating that splice inhibition strongly reduces the transcription of pri-miR-122. This was found to be unique to the endogenous gene, as it was not seen when pri-miR-122 was ectopically expressed in HeLa cells. Analysis of the chromatin-associated RNA further confirmed that splice inhibition was having a strong negative effect on transcription. Transcription of pri-miR-122 is known to be terminated by Microprocessing rather than the canonical pathway, and analysis of other lnc-pri-miRNAs which are Microprocessor terminated showed that the effect of splice inhibition on transcription was not unique to pri-miR-122, but was also shared by the important oncogene pri-miR-17~92a.

The results in this thesis show that intronic or exonic miRNA hairpins do have an effect on splicing, and unexpectedly that splicing is important for efficient transcription of specific lnc-pri-miRNAs. This work shows that the relationship between Microprocessing and splicing, as well as transcription, is complex. By elucidating how these co-transcriptional processes occur on the same transcript, the understanding of how miRNA biogenesis is regulated from different genomic locations will be improved. This could further the current knowledge of how miRNA dysregulation occurs in disease and may lead to the development of new drugs to treat them.

Acknowledgements

This is the part of my thesis that has given me the most sleepless nights. I always believed I would one day do a PhD, and I would have stuck out until the end no matter how much I hated it but I've been so lucky because I've absolutely loved my PhD experience. There are a lot of people I would like to thank and I honestly believe most of these people would have no idea the kind of impact they have made over the last 4 years.

I would firstly like to thank my supervisor, Catherine Jopling, for all her support. I am so bloody glad I ended up in your lab! I've absolutely loved my project, and if I could repeat the last 3 or so years, I would definitely choose it again. The support you have given me personally was beyond what I could have hoped for. I honestly cannot thank you enough for your kindness.

To all my friends in the Gene Regulation and RNA biology group, past and present, thank you for making everyday fun! I have to thank the 'Wang Party' (Hannah Parker, Raj Ghandi, Daniel Smalley and Jialiang Lin) as you were the first friends I made and you helped me find my feet. My brother from another mother and country, Shoaib; thank you for sharing your knowledge and your biriyani. Also, my Hot Chicken Habibi (Hibal Al-Masmoum) and Dario Balacco, I miss you both but I do not miss your arguing. I can't get away with not thanking the current lab members: Ben Hall, Jonathan Whitchurch, Bismoy Mazumder, Cintia Montiero, Mitchell Masterson, Nicole Zordane, Lorenzo Pavanello, Sadaf Ashraf. I'd particularly like to thank Poppy Winlow, Kathryn Williams and Hannah Tomlin- you have been a massive support to me and I can't thank you enough and I can only hope you will end up loving your PhDs as much as I love mine.

I also need to thank the people that make the lab tick over each day: Hilary Collins, Barbara Rampersad and Trudi Gee. I'm sorry I keep having accidents, and now I have left you should have fewer accident reports forms to fill out.

Fencing has been a massive part of my life while I have been at Nottingham. I loved working with all the girls on the 1st team and I will always have good memories of our matches and how much we have

achieved. I'd like to thank my coach, Paul Sibert for believing in me even when I didn't. I'd like to thank the fencers that have become friends I know I can depend on: Josh Petch, Jacqueline Cordell, Oscar Woolnough, Emma Duxon and Ali Delboyer.

I'd like to thank the people who don't really fit into a category. I'd like to thank Ian and Andy, my security guard friends, for making my mornings that little bit happier. I'd like to thank Conrad Nieduszynski; working with you made me believe I was good enough to do a PhD. I'd also like to thank my counsellor, Dan Pitt for making me realise how strong a person I am and that I am more determined than I realised.

To Gianni, I don't think there are enough words to do my thank you justice. How can I possibly thank you? You bought us a home to live in and you are the father of our cats. You look after me when I don't even need looking after. You always buy me biscuits even when I say not to. You make me cups of tea. These last four years haven't always been easy, but I'm so glad you have been by my side.

To the person I know I can always rely on, my Mum. Leaving you to move to Nottingham was one of the scariest things I ever did and I've miss you so much! Thank you for always believing in me. Thank you for telling me how proud of me you are. Thank you for encouraging me to do this.

I'd also like to thank the very person who made me want to be a scientist in the first place. When you died of cancer 17 years ago, I made a promise to become a scientist, to cure cancer and to make sure no child had to watch their Dad die the way I did. That promise is the reason I undertook this PhD, and hopefully one day, the work in this thesis may be used to develop that new cure for cancer, and save someone's life.

TABLE OF CONTENTS

<u>ABSTRACT</u>	<u>II</u>
<u>ACKNOWLEDGEMENTS</u>	<u>IV</u>
<u>LIST OF ABBREVIATIONS.....</u>	<u>XVI</u>
<u>CHAPTER 1: INTRODUCTION.....</u>	<u>1</u>
51.1: THE CENTRAL DOGMA OF MOLECULAR BIOLOGY	1
1.2: GENE EXPRESSION	3
1.2.1: CHROMATIN STRUCTURE	3
1.2.2: TRANSCRIPTION	4
1.2.3: CO-TRANSCRIPTIONAL EVENTS	5
1.2.3.1: POLYADENYLATION	6
1.2.3.2: SPLICING	7
1.2.3.3: CAPPING	12
1.2.4: TRANSLATION	13
1.2.5: mRNA TURNOVER	13
1.2.5.1 DEADENYLATION	15
1.2.5.2 DECAPPING	16
1.3: TYPES OF RNA	16
1.4: siRNA AND miRNA	18
1.5: GENOMIC LOCATIONS OF miRNA	18
1.6: miRNA	19
1.6.1: miRNA BIOGENESIS	20
1.6.2: TRANSCRIPTION OF miRNA	21
1.6.3: MICROPROCESSOR COMPLEX	22
1.6.3.1: HAIRPIN RECOGNITION BY THE MICROPROCESSOR	22
1.6.3.2: MODIFICATION OF DGCR8	24
1.6.3.3: MICROPROCESSOR STOICHIOMETRY	25
1.6.3.4: CO-TRANSCRIPTIONAL MICROPROCESSING	25
1.6.3.6: MICROPROCESSOR EFFICIENCY	26
1.6.3.7: DROSHA LOCALISATION	27
1.6.4: NUCLEAR EXPORT	27
1.6.5: DICER	28
1.6.7: RNA-INDUCED SILENCING COMPLEX (RISC) FUNCTION	28
1.6.7.1: RISC LOADING	29
1.6.7.2: TARGET RECOGNITION	29
1.6.7.3: RISC-MEDIATED GENE SILENCING	30
1.7: REGULATION OF miRNA BIOGENESIS	31
1.7.1: m6A METHYLATION	31
1.7.2: COFACTORS	31
1.8: miRNA TURNOVER	33
1.9: SPLICING AND miRNA BIOGENESIS	34
1.10: NON-CANONICAL PROCESSING OF miRNAs	35
1.11: NON-CANONICAL FUNCTIONS OF THE MICROPROCESSOR	37
1.12: miRNA IN LONG NON-CODING RNA	38
1.13.1: miR-122	38
1.13.2: miR-17~92	39
1.13.3: miR-23a~27a~24-2	40
1.13.4: miR-29a~29b-1	41
1.14: AIMS AND OBJECTIVES	42
<u>CHAPTER 2: METHODS</u>	<u>43</u>

2.1: GENERAL REAGENTS	43
2.2: TISSUE CULTURE	45
2.2.1: TISSUE CULTURE REAGENTS	45
2.2.2: CELL LINES	46
2.2.2.1 HeLa	46
2.2.2.2 HUH7, HUH 7.5 DERIVED CELLS	46
2.2.3 CELL MAINTENANCE	46
2.2.4 CRYOPRESERVATION	46
2.2.5: CELL THAWING	46
2.3: BACTERIAL METHODS	47
2.3.1: REAGENTS AND SOLUTIONS	47
2.3.2: BACTERIAL STRAINS	47
2.3.3: PREPARATION OF COMPETENT CELLS	47
2.3.4: TRANSFORMATION	47
2.3.5: MINIPREP	47
2.3.6: MAXIPREP	48
2.4: PLASMID TRANSFECTION	48
2.5: siRNA KNOCKDOWN	48
2.6: MORPHOLINO TRANSFECTION	48
2.7: RNA ISOLATION	49
2.8: DNASE TREATMENT	49
2.9: DETERMINATION OF NUCLEIC ACID CONCENTRATION	50
2.10: MRNA- REVERSE TRANSCRIPTION	50
2.11: MRNA- QPCR	50
2.12: miRNA- REVERSE TRANSCRIPTION	50
2.13: miRNA- QPCR	51
2.14: ANALYSIS OF QPCR DATA	51
2.15: NORTHERN BLOT FOR MRNA	52
2.15.1: SOLUTIONS	52
2.15.2 RNA AGAROSE GEL ELECTROPHORESIS	52
2.15.3: BLOTTING	52
2.15.4: CROSSLINKING AND HYBRIDISATION	53
2.15.5: PROBE PREPARATION	53
2.15.6: STRIPPING	54
2.16 NORTHERN BLOTTING FOR MATURE MIRNA	54
2.16.1: SOLUTIONS	54
2.16.2: miRNA AGAROSE GEL ELECTROPHORESIS	54
2.16.3 TRANSFER AND HYBRIDISATION	55
2.16.3: STRIPPING	55
2.16.4: PROBE PREPARATION	55
2.17: SPLICE INHIBITION	56
2.18: ALARMARBLUE CELL VIABILITY ASSAY	56
2.19: CHROMATIN-ASSOCIATED RNA EXTRACTION	56
2.20: LUCIFERASE ASSAY	56
2.21: SDS-PAGE	57
2.22: BRADFORD ASSAY	57
2.23: WESTERN BLOTTING	58
2.24: DNA AGAROSE GEL ELECTROPHORESIS	58
2.25: GEL EXTRACTION	58

CHAPTER 3: INTRONIC AND EXONIC MIR-122 EXPRESSION IN HELA 59

3.1: INTRODUCTION	59
3.2: MIR-122 EXPRESSION FROM INTRONIC AND EXONIC LOCATIONS OF B-GLOBIN	61
3.3: MIR-122 FUNCTION WHEN EXPRESSED FROM INTRONIC OR EXONIC LOCATIONS	62
3.4: INTRONIC AND EXONIC miRNA BIOGENESIS AND SPLICING COMPETITION	65

3.5: INTRONIC AND EXONIC miRNA BIOGENESIS AND MICROPROCESSOR KNOCKDOWN ON SPLICING COMPETITION	71
3.6: DROSHA CLEAVAGE EFFICIENCY	79
3.7: DISCUSSION	81
CHAPTER 4: THE EFFECT OF SPLICE INHIBITION ON PRI-MIR-122.....	84
4.1: INTRODUCTION	84
4.2: PLA B OPTIMISATION IN HELA AND HUH7 CELLS	85
4.3: THE EFFECT OF SPLICE INHIBITION AND MICROPROCESSOR KNOCKDOWN ON ENDOGENOUS PRI-MIR-122	92
4.4: SF3B1 siRNA TREATMENT HAS A SIMILAR EFFECT ON ENDOGENOUS PRI-MIR-122 AS PLA B	102
4.5: THE EFFECT OF U1 AND U2 MORPHOLINOS ON ENDOGENOUS PRI-MIR-122	105
4.6: DISCUSSION	106
CHAPTER 5: THE EFFECT OF SPLICE INHIBITION ON CHROMATIN-ASSOCIATED ENDOGENOUS PRI-MIR-122	109
5.1: INTRODUCTION	109
5.2: CHROMATIN-ASSOCIATED RNA	109
5.3: THE EFFECTS OF SPLICE INHIBITION ON CHROMATIN-ASSOCIATED PRI-MIR-122	111
5.4: THE EFFECT OF CONFLUENCY ON SPLICE INHIBITION FOR CHROMATIN-ASSOCIATED RNA	113
5.5: SPLICE INHIBITION DECREASES THE TRANSCRIPTION OF SOME, BUT NOT ALL, LNC-PRI-MIRNAS	116
5.6: THE EFFECTS OF MICROPROCESSOR KNOCKDOWN AND SPLICE INHIBITION ON CHROMATIN-ASSOCIATED RNA	120
5.7: THE EFFECT OF SPLICE INHIBITION ON PRI-MIR-122 IN CRISPR-CAS9 KNOCKOUT CELLS	126
5.8: DISCUSSION	132
CHAPTER 6: DISCUSSION.....	135
6.1: THE EFFECT OF GENOMIC LOCATION ON MIRNA BIOGENESIS	135
6.2: SPLICE INHIBITION REDUCES TRANSCRIPTION EFFICIENCY FOR ENDOGENOUS PRI-MIR-122	139
6.3: THE EFFICIENCY OF TRANSCRIPTION FOR LNC-PRI-MIRNAS	141
6.4: PRI-MIR-122 IS RAPIDLY DEGRADED	144
6.5: THE CO-TRANSCRIPTIONAL mRNA PROCESSING FACTORS AND THEIR INTERACTIONS	145
6.6: FINAL REMARKS	146
REFERENCES.....	147

TABLE OF FIGURES

Figure 1.1: The sequence hypothesis.....	1
Figure 1.2: The central dogma of molecular biology..	2
Figure 1.3: The current model of the central dogma	2
Figure 1.4: A diagram of gene expression	3
Figure 1.5: An overview of transcription..	5
Figure 1.6: A diagram of the conserved polyadenylation sequences.....	6
Figure 1.7: The conserved sequences required for splicing.....	7
Figure 1.8: Complex E of the spliceosome.	8
Figure 1.9: The formation of the prespliceosome.	8
Figure 1.10: The formation of a spliceosome with catalytic potential.	9
Figure 1.11: The formation of the catalytically active spliceosome complex B*.	10
Figure 1.12: The first transesterification reaction of the spliceosome.	10
Figure 1.13: The activation of the complex C spliceosome.....	11
Figure 1.14: The second transesterification reaction creating a spliced mRNA.....	12
Figure 1.15: The closed circle formation of mRNA	14
Figure 1.16: An overview of the deadenylation pathways	15
Figure 1.17: A schematic of miRNA biogenesis.....	21
Figure 1.18: A comparison of Drosha, DGCR8 and Dicer domains.....	22
Figure 1.19: A structural representation of a miRNA hairpin.....	23
Figure 1.20: A schematic of pri-miR-122.....	38
Figure 1.21: The order of the hairpins in pri-miR-17~92a.....	40
Figure 1.22: The order of the hairpins in pri-miR-23a~27a~24-2.....	40
Figure 1.23: The order of miR-29 family hairpins.....	41
Figure 2.1: A diagram of the northern blotting tower.....	53
Figure 3.1: Schematic of miR-122 reporter constructs.....	60
Figure 3.2: The miR-122 hairpin and the Drosha-cleavage mutant.....	60
Figure 3.3: Mature miR-122 expression is increased from a genic location compared to a lncRNA	61
Figure 3.4: Northern blot analysis shows that mature miR-122 expression increases when located within a genic region compared to a lncRNA.....	62
Figure 3.5: Schematic of firefly luciferase reporter.....	63
Figure 3.6: Repression of a luciferase reporter correlates with expression levels of miR-122 from different locations of a β -globin reporter.....	63

Figure 3.7: Repression of a luciferase reporter occurs with more than 0.1 μ M of pre-miR-122 oligo.....	64
Figure 3.8: Repression of a luciferase reporter with known concentrations of pre-miR-122 oligo.....	65
Figure 3.9: Schematic of β -globin mRNA qPCR primer binding sites.....	66
Figure 3.10: Intronic and exonic miR-122 expression from a β -globin plasmid affects host transcript splicing.....	66
Figure 3.11: A schematic showing the binding location of the northern blot probe to the β -globin transcript.....	67
Figure 3.12: The level of spliced β -globin transcripts are affected by the expression of pre-miR-122.....	68
Figure 3.13: Multiple northern blots show that the level of spliced β -globin transcripts is affected by the expression of pre-miR-122.....	69
Figure 3.14: Combined northern blot analysis shows that miR-122 expression from an exon leads to a decrease in the levels of host transcript splicing.....	70
Figure 3.15: Intronic and exonic miR-122 expression may increase the levels of unspliced host transcript.....	71
Figure 3.16: Mature miR-122 expression from β -globin reporters containing pre-miR-122 following Microprocessor knockdown.....	72
Figure 3.17: Mature miR-122 expression when located within a coding region following Microprocessor knockdown.....	73
Figure 3.18: Intronic and exonic miR-122 expression may affect host transcript splicing following Microprocessor knockdown.....	74
Figure 3.19: MiR-122 expression from an exon or intron may affect host transcript splicing following Microprocessor knockdown.....	75
Figure 3.20: The level of spliced β -globin transcripts are affected by the expression of pre-miR-122 following Microprocessor knockdown.....	76
Figure 3.21: A graph visualising the northern blot quantification of spliced β -globin.....	77
Figure 3.22: Intronic and exonic miR-122 expression may affect host transcript splicing following Microprocessor knockdown.....	78
Figure 3.23: Intronic and exonic miR-122 expression may affect host transcript splicing following Microprocessor knockdown when compared to DMSO treated cells.....	79
Figure 3.24: Schematic of qPCR primer bind sites used to measure Drosha cleavage efficiency.....	80
Figure 3.25: DGCR8 knockdown affects Drosha cleavage efficiency.....	80

Figure 4.1: The structure of pla B.....	84
Figure 4.2: Schematic of pri-miR-122.....	85
Figure 4.3: A schematic of the qPCR binding sites in pri-miR-122.....	85
Figure 4.4: Splice inhibition using pla B reduces the levels of spliced pri-miR-122.....	86
Figure 4.5: Splice inhibition using pla B reduces the levels of endogenous but not overexpressed unspliced pri-miR-122.....	87
Figure 4.6: Splice inhibition using pla B reduces the levels of spliced pri-miR-122 within 1 hr of treatment.....	88
Figure 4.7: Splice inhibition leads to an increase in unspliced pri-miR-122 when overexpressed in HeLa cells.....	89
Figure 4.8: Mature miR-122 after splice inhibition.....	90
Figure 4.9: Short treatments of pladienolide B are not cytotoxic to HeLa cells.....	91
Figure 4.10: Short treatments of pladienolide B are not cytotoxic to Huh7 cells.....	92
Figure 4.11: DGCR8 knockdown leads to a reduction in DGCR8 and Drosha protein expression.....	93
Figure 4.12: A schematic of the qPCR binding sites in pri-miR-122.....	93
Figure 4.13: Splice inhibition or Microprocessor knockdown affects the levels of endogenous unspliced pri-miR-122.....	94
Figure 4.14: Splice inhibition in conjunction with Microprocessor knockdown may affect the levels of endogenous unspliced pri-miR-122.....	95
Figure 4.15: Splice inhibition or Microprocessor knockdown affects the levels of endogenous spliced pri-miR-122.....	96
Figure 4.16: Splice inhibition in conjunction with Microprocessor knockdown has little effect on the levels of endogenous spliced pri-miR-122.....	97
Figure 4.17: Microprocessor knockdown leads to a decrease in miR-122 processing	98
Figure 4.18: Splice inhibition with Drosha knockdown affects the levels of endogenous pri-miR-122 hairpin compared to pla B treated siRNA control.....	99
Figure 4.19: Splice inhibition leads to an increase in pri-miR-122 transcriptional read through	100

Figure 4.20: Splice inhibition has no additional effect on transcriptional termination efficiency of pri-miR-122 with Microprocessor knockdown.....	101
Figure 4.21: Splice inhibition using pla B and Microprocessor knockdown has little effect on the levels of endogenous mature miR-122.....	102
Figure 4.22: Splice inhibition using SF3b1 siRNA reduces the levels of endogenous spliced and unspliced pri-miR-122.....	103
Figure 4.23: Outlier removal shows that splice inhibition using SF3b1 siRNA reduces the levels of different unspliced lnc-pri-miRNAs.....	104
Figure 4.24: The effect of splice inhibition on mature miR-122 using SF3b1 siRNA.....	105
Figure 4.25: U1 and U2 AMOs lead to a reduction in spliced and unspliced pri-miR-122.....	106
Figure 5.1: Western blot analysis of fractions obtained following the chromatin-associated RNA extraction protocol.....	110
Figure 5.2: A schematic of the qPCR binding sites in pri-miR-122.....	110
Figure 5.3: Chromatin-associated RNA extraction protocol separates nucleoplasmic RNA from chromatin-associated RNA.....	111
Figure 5.4: Splice inhibition reduces the levels of Chromatin-associated pri-miR-122.....	112
Figure 5.5: Confluency has little effect on chromatin-associated unspliced pri-miR-122 following splice inhibition.....	113
Figure 5.6: Confluency has little effect on chromatin-associated spliced pri-miR-122 following splice inhibition.....	114
Figure 5.7: Confluency may affect Drosha cleavage efficiency following splice inhibition.....	115
Figure 5.8: Confluency has little effect on chromatin-associated pri-miR-122 3' read through following splice inhibition.....	116
Figure 5.9: Splice inhibition leads to a reduction of some unspliced lncRNA containing miRNA hairpins.....	117
Figure 5.10: Confluency has little effect on chromatin-associated unspliced pri-miR-17~92a following splice inhibition.....	118
Figure 5.11: Confluency has little effect on chromatin-associated unspliced pri-miR-23a~24-2~27a following splice inhibition.....	119
Figure 5.12: Confluency has little effect on chromatin-associated unspliced pri-miR-29a~29b-1 following splice inhibition.....	120

Figure 5.13: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated spliced pri-miR-122.....	121
Figure 5.14: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated unspliced pri-miR-122.....	121
Figure 5.15: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated pri-miR-122 hairpin.....	122
Figure 5.16: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated pri-miR-122 transcriptional read through.....	123
Figure 5.17: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated unspliced pri-miR-17~92a.....	124
Figure 5.18: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated unspliced pri-miR-23a~27a~24-2.....	125
Figure 5.19: Splice inhibition using pla B has no effects on the levels of endogenous chromatin-associated unspliced pri-miR-29a~29b-1.....	126
Figure 5.20: Splice inhibition using pla affects the levels of endogenous chromatin-associated spliced pri-miR-122 in Huh 7.5 knock out cells.....	127
Figure 5.21: Splice inhibition using pla affects the levels of endogenous chromatin-associated unspliced pri-miR-122 in Huh 7.5 knock out cells.....	128
Figure 5.22: Splice inhibition using pla B affects Drosha cleavage of the miR-122 hairpin in chromatin-associated pri-miR-122 in Huh 7.5 knock out cells.....	129
Figure 5.23: Splice inhibition using pla effects transcriptional termination of chromatin-associated pri-miR-122 in Huh 7.5 knock out cells.....	130
Figure 5.24: Splice inhibition using pla effects transcriptional termination of chromatin-associated unspliced pri-miR-17~92~92a in Huh 7.5 knock out cells.....	130
Figure 5.25: Splice inhibition using pla effects transcriptional termination of chromatin-associated unspliced pri-miR-23a~24-2~27a in Huh 7.5 knock out cells.....	131
Figure 5.26: Splice inhibition using pla B effects transcriptional termination of chromatin-associated unspliced pri-miR-29a~29b-1 in Huh 7.5 knock out cells.....	132

Figure 6.1: Schematic of miR-122 β -globin constructs.....	136
Figure 6.2: Schematic of miR-155 reporter constructs.....	138
Figure 6.3: Schematic of miR-155 and miR-122 within the same reporter constructs.....	138

TABLE OF TABLES

Table 1: Plasmids	44
Table 2: Antibodies	45
Table 3: siRNA sequences	45
Table 4: A table of qPCR primers used- update	51
Table 5: SDS-PAGE	57
Table 6: The effect of splice inhibition of pri-miR-122	140
Table 7: The effect of Microprocessor knockdown or splice inhibition of Inc-pri-miRNAs	142
Table 8: The relative expression of total and chromatin-associated pri-miR-122 transcripts following splice inhibition and Microprocessor knockdown	144

List of Abbreviations

AGO	Argonaute
AMO	Morpholino
APS	Ammonium persulphate
ATP	Adenosine 5'- triphosphate
BP	Base pair
BSA	Bovine serum albumin
CBC	Cap binding complex
CBP	Cap binding protein
cDNA	Complementary deoxyribonucleic acid
CDS	Coding sequence
<i>C.elegans</i>	<i>Caenorhabditis elegans</i>
CFIm	Cleavage factor I
CFIIm	Cleavage factor II
CIM	CCR4-interacting motif
circRNA	Circular RNA
CO ₂	Carbon dioxide
CPSF	Cleavage and polyadenylation factor
CstF	Cleavage stimulation factor
CTD	C-terminal domain
dATP	Deoxyadenosine 5'-triphosphate
dCTP	Deoxycytidine 5'-triphosphate
DGCR8	DiGeorge critical region 8
DMEM	Dulbecco's modified eagles medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide 5-triphosphate
DSE	Downstream element
dsRNA	Double stranded RNA
DTT	Dithiothreitol
DUF	Domain of unknown function
ECL	Enhanced chemiluminescence
EDTA	Ethylene diamine tetraacetic acid
eIF	Eukaryotic initiation factor
Exp	Exportin
FUS	Fused in sarcoma/translocated in liposarcoma
HCV	Hepatitis C virus

HeLa	Human cervical epithelial carcinoma cells
IRES	Internal ribosome entry site
Kb	Kilobase
kDa	Kilodalton
LB	Lysogeny broth
LLR	Leucine-rich repeat
LncRNA	Long non-coding RNA
Lnc-miRNA	Long non-coding miRNA
m ⁷ G	7-methylguanosine
METTL3	Methyltransferase-like 3
miRISC	miRNA- RNA induced silencing complex
miRNA	MiRNA
MOV10	Moloney leukemia virus 10
MP	Microprocessor
mRNA	Messenger ribonucleic acid
mRNP	Messenger ribonucleoprotein
NC	Non-coding
NCR	Non-coding region
NEAA	Non-essential amino acid
NES	Nuclear exportation signal
NLS	Nuclear localisation signal
NRO	Nuclear run on
NT	Nucleotide
PACT	Protein kinase RNA activator
PAGE	Polyacrylamide gel electrophoresis
PAN2	PABP-dependent polyA nuclease 2
PAN 3	PABP-dependent polyA nuclease 3
PAP	PolyA polymerase
PAS	Polyadenylation signal
PAZ	PIWI, AGO, and zwillie
P Bodies	Processing bodies
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
Pla B	Pladienolide B
Pol II	RNA polymerase ii
Pre-miR	Pre miRNA
P Rich	Proline rich
Pri-miR	Primary micro ribonucleic acid

RF	Release factor
Rhed	RNA-binding heme domain
RISC	RNA-induced silencing complex
RITS	RNA induced transcriptional silencing
RT	Room temperature
RNA	Ribonucleic acid
RNase	Ribonucleic acid hydrolase
RNasin	Ribonuclease inhibitor
RPA	RNA protection assays
RRM	RNA recognition motif
rRNA	Ribosomal RNA
RS-rich	Arginine/serine rich
scaRNA	Small cajal body-specific RNA
siRISC	siRNA- RNA induced silencing complex
SDS	Sodium dodecyl sulphate
siRNA	Small interfering ribonucleic acid
SOMiR	Splice site overlapping miRNA
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
snRNP	Small nuclear ribonucleoprotein
SS	Single stranded
SSIII	Superscript III
SSC	Saline sodium citrate
ssRNA	Single stranded RNA
stRNA	Short temporal RNA
SUMO	Small ubiquitin-like modifier
TAE	Tris acetic acid EDTA buffer
TBS	Tris buffered saline
TBST	Tris Buffered saline tween
TEMED	Tetramethylethylenediamine
TN	Trans-dominant negative
TRBP	TAR RNA binding protein
tRNA	Transfer RNA
TSS	Transcriptional start site
TU	Transcriptional unit
UTR	Untranslated region
UV	Ultraviolet
Xist	X chromosome specific transcript

Chapter 1: Introduction

51.1: The central dogma of molecular biology

The central dogma of molecular biology originated in 1958 when much of what we know now was not even conceivable. 12 years later, according to Francis Crick who developed the term central dogma, molecular biology was 'a boundless optimism that the basic concepts involved were rather simple and probably much the same in all living things'. What Crick is alluding to is that the 'Sequence hypothesis,' which hypothesised that information on how to create a protein is passed on linearly with DNA coding only for RNA, and RNA a template for creating protein (**Figure 1.1**)¹.

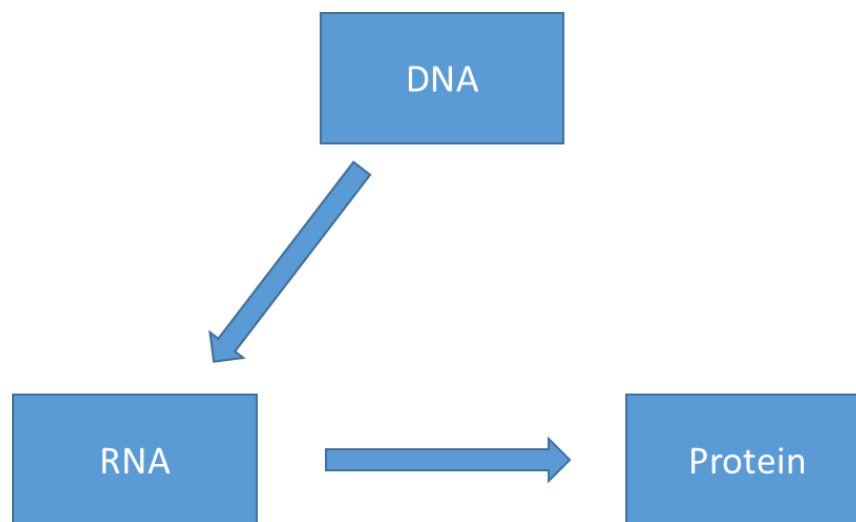


Figure 1.1: The sequence hypothesis

The central dogma takes a more complex view on the transfer of information between DNA, RNA and protein (**Figure 1.2**). The dogma developed the idea that three classes of information transfer can occur; general, special and unknown. General transfers are information transfers known to occur, such as DNA can create DNA and RNA, and RNA can create protein. Special transfers were thought to occur in certain circumstances, such as RNA creating RNA and DNA, and DNA directly creating protein. Thought to not occur, the process of protein creating protein, DNA and RNA make up the unknown transfers group.

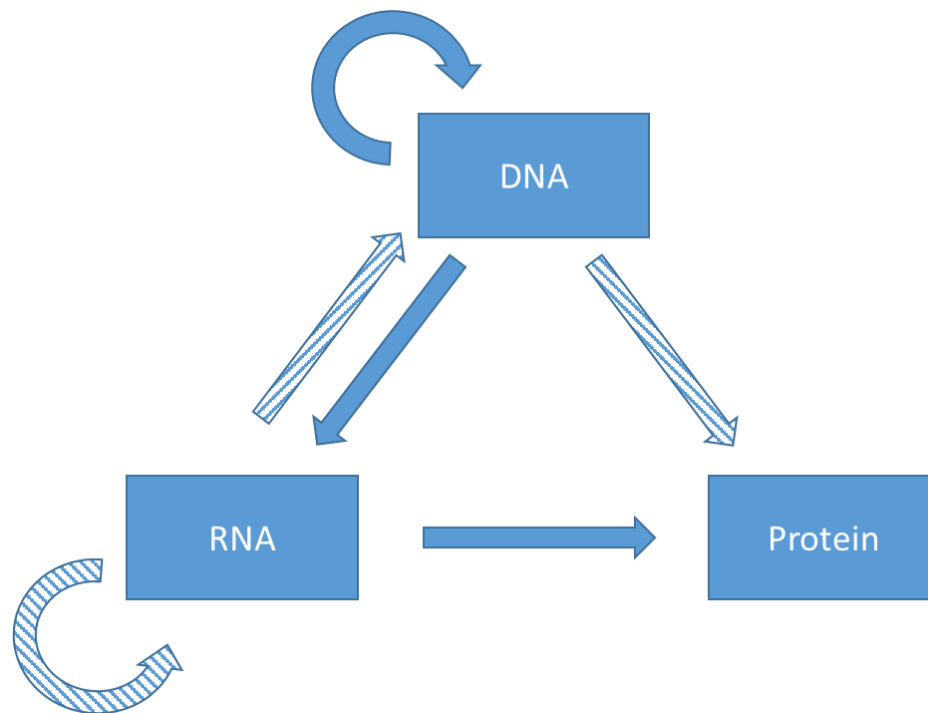


Figure 1.2: The central dogma of molecular biology. The block arrows represent general transfers. The striped arrows represents special transfers.

The overview of the central dogma of molecular biology will seem very simplistic to modern-day researchers. The dogma was updated in 1970 when the first papers on reverse transcription were published^{2,3} and both reverse transcription and RNA replication are now well characterised processes. It has also been indicated that protein synthesis, the third special transfer, could occur directly from DNA, although there is not much evidence for this⁴. The first iteration of the central dogma stated: there is no information transfer from protein to nucleic acid¹. As yet, there is no evidence for reverse translation and that this process may be too complex to reverse engineer (**Figure 1.3**)⁵.



Figure 1.3: The current model of the central dogma

Other transfers of information have since been discovered and are difficult to place in the central dogma. These include post-translational modifications of proteins, intein-mediated protein splicing, epigenetic modifications, prions, and possibly RNA editing⁵.

1.2: Gene expression

The concept of a 'gene' is hard to define. Initially, a gene was thought to be a region of DNA that encoded protein. The definition of a gene has become muddled of late, with a single gene being able to create a multitude of products, which are not necessarily proteins. Gene expression is a term used to describe the process of creating a product from a gene beginning with transcription- the process of creating an RNA transcript from a DNA code. The RNA transcript can be modified in a variety of ways to affect gene expression, such as polyadenylation and capping, while an RNA transcript can also be targeted for degradation, and the initial transcription can be controlled at the chromatin level (**Figure 1.4**).

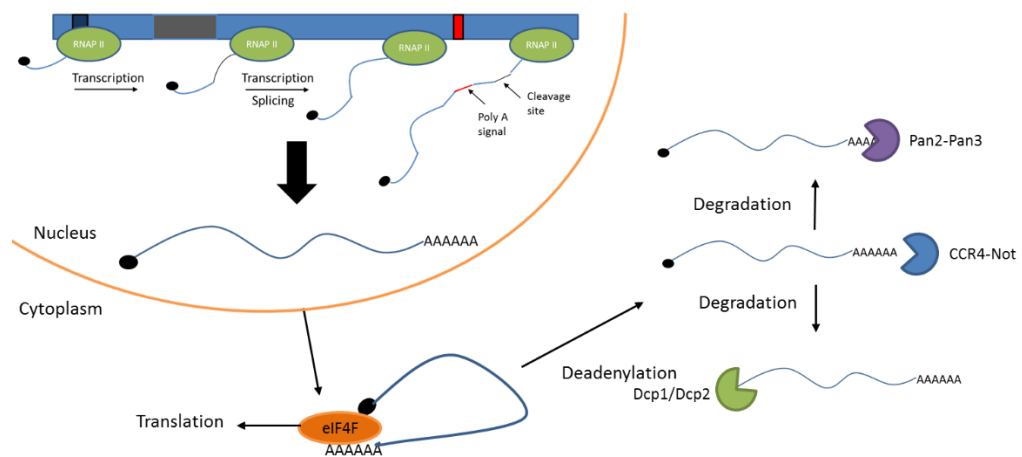


Figure 1.4: A diagram of gene expression

1.2.1: Chromatin structure

In eukaryotes, DNA is wrapped around histone proteins to produce a macromolecule complex called chromatin. Chromatin functions as a way of tightly packaging DNA to protect it from DNA damage and to control gene expression and DNA replication. The structure of the chromatin can vary throughout the cell cycle. During interphase, where the cell is undergoing DNA replication to double the amount of chromosomes prior to mitosis, the chromatin takes on a loosely structured complex to allow DNA polymerases to replicate the DNA.

In 1964, a paper was published investigating the role of histones in transcription. This work showed that histone proteins could be acetylated

and methylated after protein synthesis, and that the levels of acetylation directly affected RNA polymerase activity⁶. The authors conclude that the levels of acetylation may be altered to allow for more or less gene expression at different times and different loci.

It is now clear that histone modifications play a vital role in gene expression, and the study of histone modifications is now a major part of the widely researched field of epigenetics. Histone modifications have been shown not only to regulate chromatin structure, but to recruit proteins to the chromatin. A range of histone modifications have now been identified such as phosphorylation, ubiquitination, sumoylation, and not forgetting methylation and acetylation. New modifications are still being identified with the twelfth and thirteenth histone modifications, crotonylation and hydroxylation, identified in 2011⁷.

1.2.2: Transcription

Transcription is the process of a DNA sequence being copied into RNA by a RNA polymerase enzyme (**Figure 1.5**). There are three types of nuclear RNA polymerases in animals. RNA polymerase I transcribes ribosomal RNA (rRNA), RNA polymerase III transcribes small non-coding RNAs like transfer RNA (tRNA), and RNA polymerase II transcribes protein coding genes and will be the focus for this section. Transcription factors are necessary for RNA polymerases to function, which aid promoter recognition, RNA polymerase recruitment, DNA unwinding and recognition of the transcriptional start site (TSS). The first stage in transcription is the binding of the polymerase to the promoter region via the transcription factors. Once the polymerase is recruited, the promoter region is 'opened' and allows the polymerase to travel along the DNA until the transcriptional start site (TSS) is located⁸⁻¹⁰.

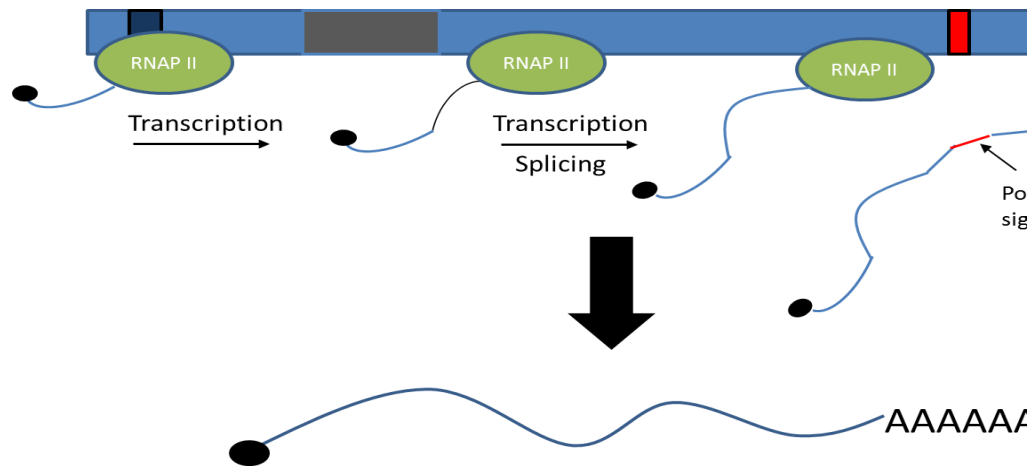


Figure 1.5: An overview of transcription. The dark blue box represents the TSS, the blue boxes represent exons with the grey box indicating an intron. The red box is the location of the PAS. In the mRNA transcript, the black circle represents the cap.

The elongation step of transcription is the process of the RNA polymerase travelling along the template strand of the DNA in a 3' to 5' direction and pairing the DNA bases with their complementary RNA bases. This produces an RNA transcript running from a 5' to 3' direction. During transcriptional elongation, co-transcriptional processes occur such as capping, splicing, polyadenylation and miRNA biogenesis.

Transcriptional termination is initiated by the transcription of the polyA signal. The polyA signal in the mRNA is recognised by two proteins: cleavage and polyadenylation specificity factor (CPSF) and cleavage stimulation factor (CstF). These two proteins recruit other proteins to the site leading to the mRNA being cleaved. During this process, the RNA polymerase continues to transcribe until the mRNA is cleaved and the RNA polymerase is released from the DNA. Once cleaved, a polyA tail is added to the mRNA.

1.2.3: Co-transcriptional events

The carboxy-terminal domain (CTD) is the largest domain of RNA polymerase II and consists of 52 heptad repeats in mammals with the consensus sequence: YSPTSPS. The CTD is subject to a range of post-translational modifications. The tyrosine, threonine and the serines can all be phosphorylated¹¹, while the threonine and serines can also be glycosylated¹². The prolines can switch between cis and trans forms¹³. The end of the heptad repeats contains a non-consensus repeat which may

include arginine and lysine residues- these can be methylated¹⁴ or ubiquitinated¹⁵. Calculations suggest that there could be 10^{59} combinations of post-translational modifications on the CTD¹⁶.

The CTD protrudes from one of the 12 protein subunits located near the mRNA exit site in RNA polymerase II and provides a binding site for regulatory proteins, allowing mRNA processing to be coupled to transcription¹⁷. The CTD has been shown to be involved with splicing and 3' processing¹⁸ and different segments of the CTD have different functions¹⁹. Together, this information suggests that mRNA could be processed during transcription in a variety of ways in parallel.

1.2.3.1: Polyadenylation

Polyadenylation is the process of adding a length of adenosine bases to the 3' end of a RNA transcript which begins once transcription has been terminated. Cleavage of the 3' end is coupled with polyadenylation and mediated by five different proteins: polyA polymerase (PAP), cleavage and polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF), and cleavage factors I (CFIm) and II (CFIIm).

In mammals there are three highly conserved sequence motifs involved in polyadenylation: the polyadenylation signal (PAS), the downstream element (DSE), and the cleavage site itself (**Figure 1.6**)^{20,21}. The PAS is a hexanucleotide sequence (AAUAAA) found in approximately 90 % of mRNA transcripts 10-30 nucleotides upstream of the cleavage site. The DSE is a GU-rich region downstream of the cleavage site and is generally less well conserved compared to the PAS and is located around 30 nucleotides from the cleavage site. As already indicated, the cleavage site is located between the PAS and the DSE in a window of 30 nucleotides²².

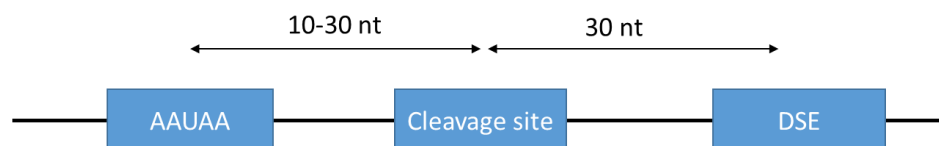


Figure 1.6: A diagram of the conserved polyadenylation sequences

CPSF recognises and binds the PAS sequence, while CstF associates with the GU-rich region and enhances the binding efficiency of CPSF^{21,23-25}. CFIm and CFIIm have been shown to cleave the 3' end of pre-mRNA in

the presence of CstF and CPSF^{22,26} The cleavage factor CFIm has been shown to recognise a UGUAA domain upstream of the PAS and has a role in stabilising the RNA-CPSF interaction^{22,27,28}.

Elongation of the polyA tail occurs in two stages. Firstly, CPSF recruits PAP and leads to the addition of the first 10-12 adenosines of the tail^{29,30}. Secondly, the short tail is recognised by PABP (PolyA binding protein) which leads to the extension of the tail up to 250 adenosines with more PABP binding to the tail during synthesis. PABP, PAP and CPSF interact to stabilise polyadenylation, however, when the tail reaches a certain length the complex is destabilised ending polyadenylation³⁰.

1.2.3.2: Splicing

The removal of introns from an mRNA sequence involves the spliceosome- a highly dynamic complex of five small nuclear RNAs (snRNA), known as U1, U2, U4, U5 and U6, which associate with proteins to form snRNPs (small nuclear RNA protein complex)³¹. As each intron is transcribed by RNA polymerase II, the assembly of a de novo spliceosome occurs, and with the average gene containing eight introns, it can take many spliceosomes to process a pre-mRNA into a mature mRNA transcript. How many spliceosomes can assemble onto a mRNA transcript is currently unknown³². While splicing is a highly conserved process, higher eukaryotes often contain degenerate splice sites and a larger variety of splicing factors which can modulate splicing, and lead to processes such as alternative splicing. Similarly, splicing may also influence other processes, such as transcription and chromatin regulation³³.

Splicing requires the presence of conserved sequences called splice sites (SS) which define the 5' and 3' intron boundaries (**Figure 1.7**). These splice sites consist of GU and AG dinucleotides^{34,35}. A polypyrimidine tract, about 18-20 nt in length helps identify the 3' SS during spliceosome assembly. At the end of the polypyrimidine tract is the branch-point sequence (BPS)^{36,37}.

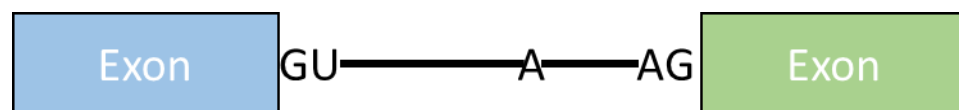


Figure 1.7: The conserved sequences required for splicing. The black line represents the intron. The A represents the adenosine which forms part of the branch-point sequence.

There are many stages to spliceosome assembly: E, A, B, B^{act}, B*, C, C*, and P. The first step in spliceosome assembly is the association of the E complex. The U1 snRNP which associates with RNA polymerase II independently of splicing binds to the 5' SS. This interaction causes splicing factor 1 (SF1) and U2 auxiliary factor (U2AF) to interact with the BPS and polypyrimidine tract (**Figure 1.8**)³².



Figure 1.8: Complex E of the spliceosome. The black line represents the intron.

The second step is the conversion of complex E to the pre-spliceosome complex A. This involves SF1 dissociating, and the U2 snRNP binding to the BPS, which causes an adenosine to bulge out of the pre-mRNA-U2 complex (**Figure 1.9**).

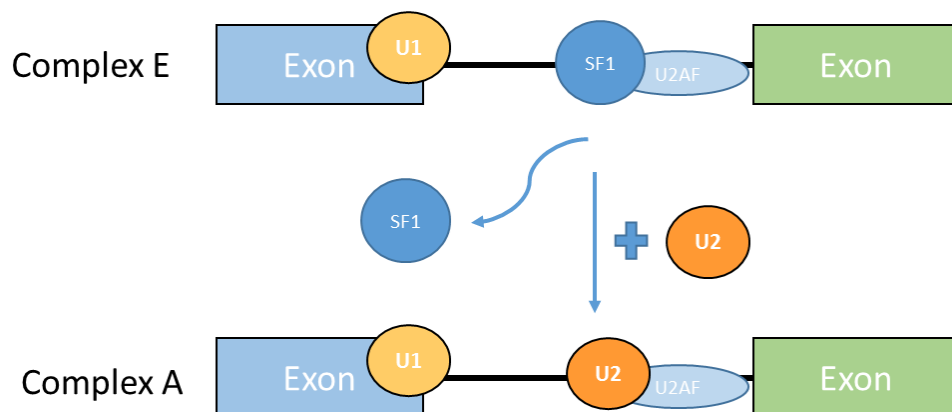


Figure 1.9: The formation of the prespliceosome.

The next stage is the recruitment of the tri-snRNP consisting of U4, U6, and U5 which results in the formation of complex B³⁸⁻⁴⁰. Changes in RNA-RNA and RNA-protein interactions, along with the recruitment of spliceosome associated factors and the release of U1 and U4 snRNP characterises the maturation of the B complex to B^{act}⁴¹. The interactions between U2 and U6 form the catalytic site of the spliceosome, although in B^{act} this is not yet functional (**Figure 1.10**).

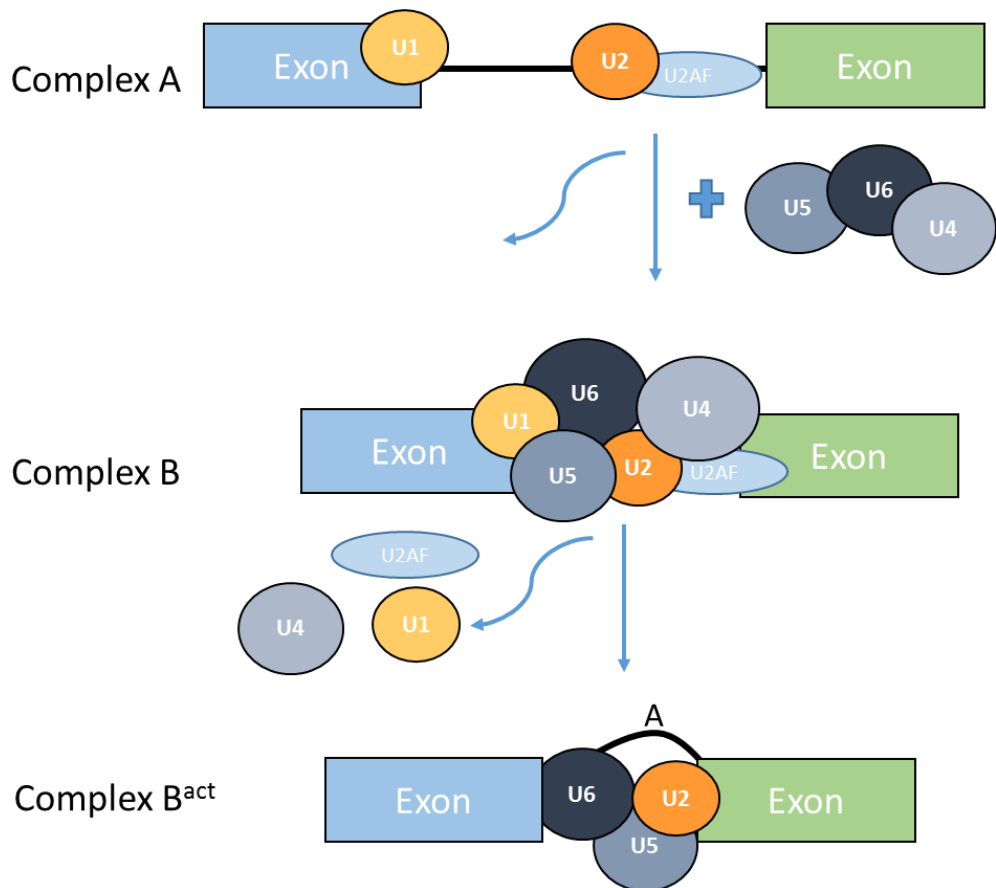


Figure 1.10: The formation of a spliceosome with catalytic potential.

UV crosslinking of the yeast B^{act} complex indicated that the SF3a and SF3b complexes of the U2 snRNP interact with the RNA transcript either side of the BPS^{42,43}. An ATPase Prp2 is responsible for the conversion of B^{act} into a catalytically active B* complex by destabilising SF3a and SF3b. This leads to the adenosine bulge formed in complex A to be exposed (**Figure 1.11**).

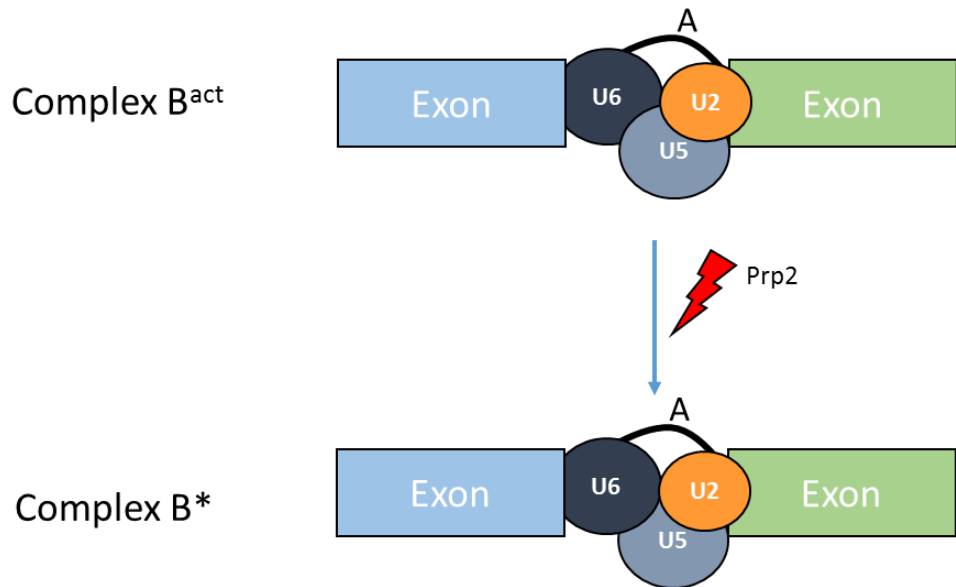


Figure 1.11: The formation of the catalytically active spliceosome complex B*.

The B* complex undertakes the first transesterification step: The 2' OH group of the adenosine of the intron is involved in the nucleophilic attack on the 5' SS resulting in cleavage and the ligation of the intron to the branch adenosine^{43,44}. This leads to the formation of an intron lariat structure (**Figure 1.12**).

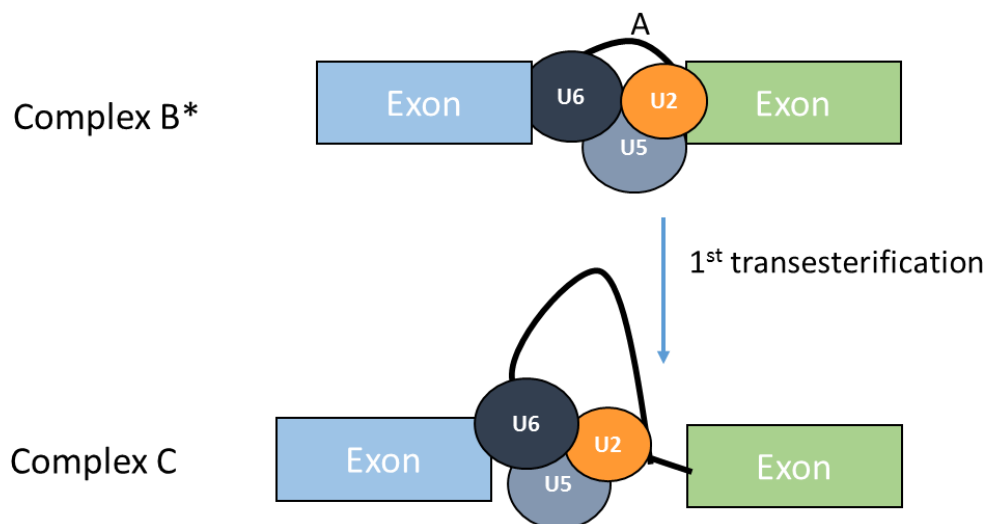


Figure 1.12: The first transesterification reaction of the spliceosome.

Complex C is formed following step I transesterification. The U5 snRNP base pairs with the 5' exon, the U6 snRNP interacts with the 5'ss, and the U2 snRNP binds to the intron lariat structure^{45,46}. The U5 snRNP protein Prp16, another ATPase, rearranges the catalytic site which leads to the

formation of the activated complex C* which is able to carry out the final transesterification step (**Figure 1.13**)^{47,48}. The active site of the C* complex binds to the first AG dinucleotide downstream of the BPS⁴⁹. This is referred to as the 'first AG rule', however, there are incidences where the AG dinucleotide is not recognised^{50,51}.

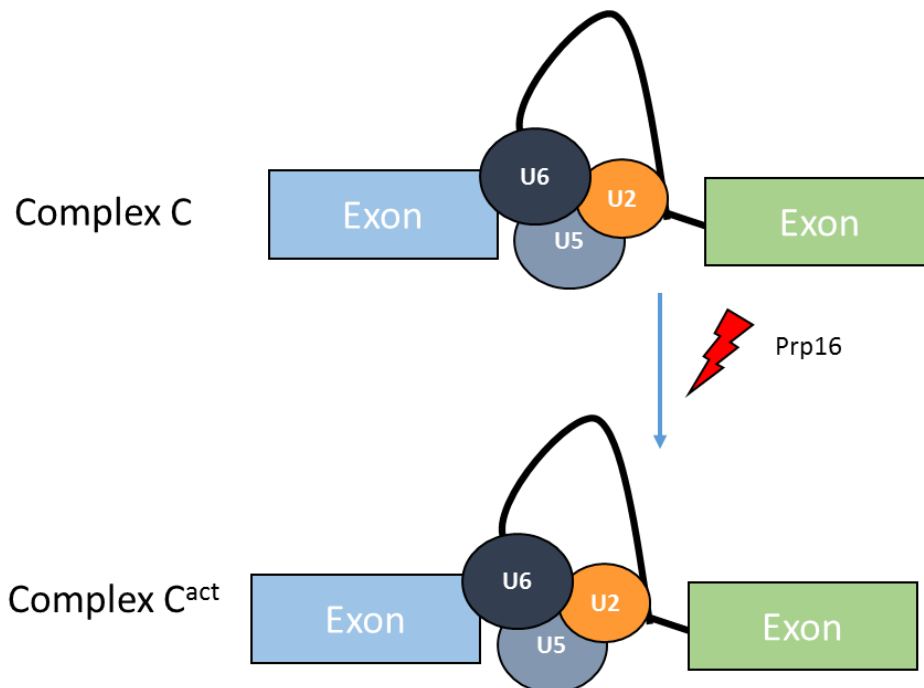


Figure 1.13: The activation of the complex C spliceosome.

In the second and last transesterification reaction, the 3' OH group of the 5' exon is attacked by the 3'SS, and leads to the ligation of the 5' and 3' exons⁵². After this stage, the spliceosome (complex P) is still associated with the ligated exon and the excised intron lariat (**Figure 1.14**). The ATPase Prp43 disassembles the complex P spliceosome into the individual components which can be recycled for future splicing events⁵³.

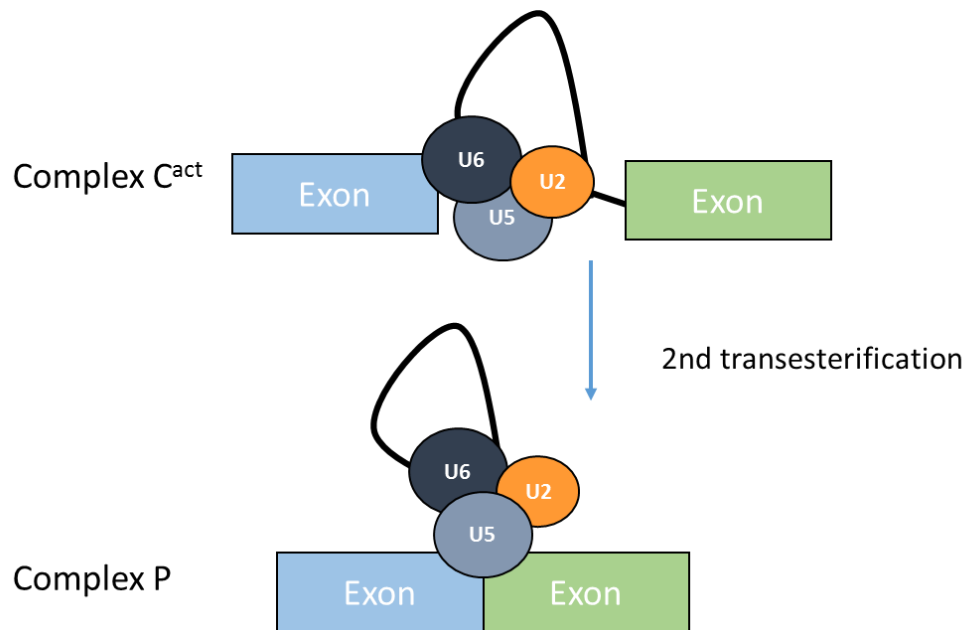


Figure 1.14: The second transesterification reaction creating a spliced mRNA.

While this describes the majority of splicing activity that can occur within an mRNA transcript, there is a second method of splicing that requires the 'minor' spliceosome as opposed to the 'major' spliceosome which has just been described. The minor spliceosome removes U12-type introns and results in less than 1 % of the splicing activity of a cell. The minor spliceosome also requires different snRNPs: U11, U12, U4atac, and U6atac. Both spliceosomes require U5⁵⁴.

1.2.3.3: Capping

A cap is added to the 5' end of mRNA during transcription by a methyltransferase. The process of capping occurs in three steps. Firstly, the γ phosphate is removed from the first nucleotide of the newly synthesised RNA transcript to create a 5'-diphosphate end. Secondly, a guanylyl-transferase reacts with GTP to produce a GMP intermediate which is added to the 5'-diphosphate connecting the guanosine to the RNA transcript via a triphosphate bridge. Thirdly, a methyl group is added to the N7 position of the guanine from an *S*-adenosylmethionine to create a 7-methylguanosine cap, or M⁷G⁵⁵⁻⁵⁸.

A cap-binding complex (CBC) recognises the 7-methylguanosine cap and has a role in pre-mRNA recognition and efficient co-transcriptional processes. The CBC consists of two proteins: CBP80 and CBP20⁵⁹. The CBC

has been shown to have a role in splicing efficiency, and is necessary for the recognition of the 5' splice site of introns proximal to the cap by the U1 snRNP⁶⁰. The CBC has also been shown to have a role in the export of RNA transcripts from the nucleus into the cytoplasm^{55,60-62}. Polyadenylation and 3' cleavage efficiency are enhanced by the CBC, which demonstrates an interaction between the 5' and 3' ends of a nascent RNA transcript⁶³. The 7-methylguanosine cap also has a role in mRNA turnover and translation.

1.2.4: Translation

Translation is the process of creating a protein from an mRNA transcript. Like transcription, translation occurs in three steps: initiation, elongation and termination. There are two methods by which initiation can occur: cap dependent and cap independent.

Cap-dependent translation initiation requires the m7G and is the process of the assembly of 80S ribosomes on the mRNA leading to the initiation codon, AUG, base pairing with the anticodon Met-tRNA^{Met}. A minimum of nine eIFs are involved, including eIF4E which binds to the m7G cap, and eIF4G which associates with PABP. The first step in translation initiation is the binding of a 43S pre-ribosomal unit to the 5' cap proximal region of mRNA, which consists of a 40S ribosome, the initiator anticodon Met-tRNA^{Met} and eIFs. The 43S pre-ribosomal unit scans the mRNA in search of an initiation codon in a 5' to 3' direction. When the initiation codon has been recognised, the eIFs are displaced and a 60S ribosome is attached to produce an 80S ribosome⁶⁴.

Cap-independent translation initiation does not require a 5' cap, but instead uses an internal ribosome entry site, or IRES and is a common method of translation initiation in viruses, including HCV^{65,66} and some cellular mRNAs. IRESs are highly structured regions in the 5' UTR containing both secondary and tertiary structures and have been grouped into 4 classes depending on their requirement the initiator anticodon Met-tRNA^{Met} and eIFs⁶⁷. Regardless of the class, the IRES (and the necessary factors) interacts with the 40S ribosome to initiate translation.

1.2.5: mRNA turnover

An mRNA transcript that has been capped, polyadenylated and spliced is not necessarily translated into a protein and may be degraded within the

cytoplasm. This process is called mRNA turnover and is another mechanism of maintaining proper gene expression. The 3' polyA tail and the 5' cap act together to protect the mRNA from degradation from exonucleases by forming a closed loop structure via the binding of the eukaryotic initiation factor (eIF) eIF4F which primarily has a role in promoting translation initiation (**Figure 1.15**). eIF4F consists of three subunits: the DEAD-box helicase eIF4A, the cap binding protein eIF4E and eIF4G. For the mRNA to be degraded, either the polyA tail or the cap need to be exposed to exonucleases, or the mRNA is targeted internally by endonucleases⁶⁸.

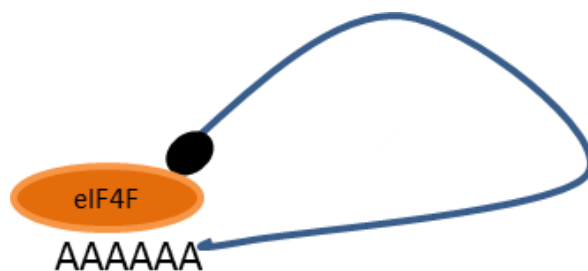


Figure 1.15: The closed circle formation of mRNA

The exosome is involved in the nuclear degradation of a variety of mRNAs, including unstable products of transcription, and transcripts that have been incorrectly processed by the co-transcriptional machinery and degrades mRNA in a 3'-5' direction. The exosome is a barrel-shaped complex consisting of 10 or 11 ribonucleolytic subunits and mRNAs are targeted to the exosome by exosome-specificity factors which are known to recognise elements representing unprocessed or unstable transcripts, such as AU-repeats in the 3' UTR⁶⁹.

mRNA can also be degraded in the nucleus by the 5' -3' exonuclease Xrn2 following co-transcriptional cleavage of the transcript⁷⁰. Errors in the maturation of mRNA leads to the mRNA transcript being co-transcriptionally degraded by Xrn2. Interestingly, splicing inhibition by Spliceostatin A has been shown to lead to an increase in Xrn2-mediated degradation⁷¹. Knockdown of Xrn2 has also been shown to lead to an increase in transcriptional pausing⁷².

1.2.5.1 Deadenylation

In eukaryotes, the majority of mRNA decay begins with the removal of the polyA tail via a process called deadenylation (**Figure 1.16**). The deadenylation step occurs in two phases. The initial shortening of the polyA tail is mediated by the PAN2-PAN3 deadenylation complex, where the tail is shortened to around 10-25 nucleotides⁷³. At this stage, the shortened, tail can still be readenylated. Once the tail has been shortened, the polyA tail is further degraded by the CCR4-NOT complex^{70,72-73}. Ccr4 is one of two catalytic subunits and contains a Mg^{2+} dependent polyA specific 3' to 5' exonuclease^{68,72}. Ccr4 contains a leucine-rich repeat (LRR) which interacts with the second catalytic subunit, Caf1. Caf1 is also a 3' to 5' exonuclease.

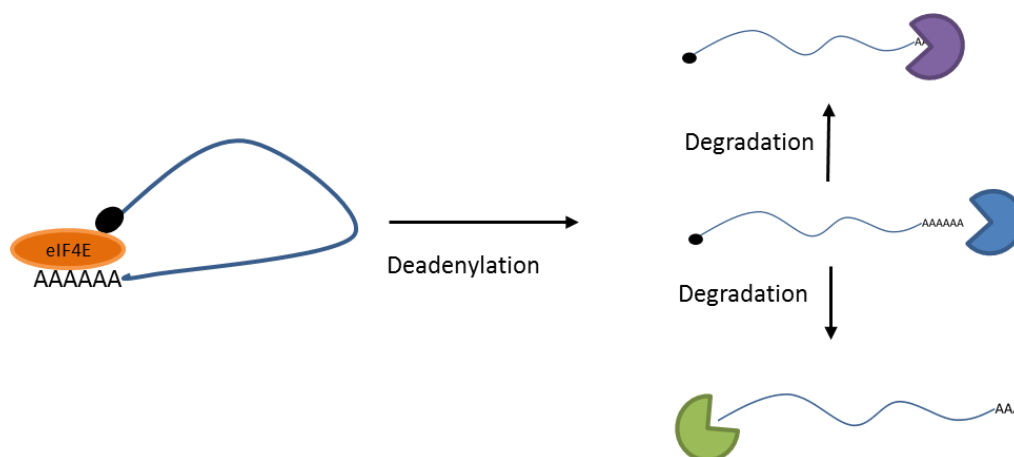


Figure 1.16: An overview of the deadenylation pathways

Confusingly, both Ccr4 and Caf1 have two paralogs, and each can be referred to by different names. Ccr4 can be referred to as CNOT6/Ccr4a or CNOT6L/Ccr4b while Caf1 may be referred to by its paralogs CNOT7/Caf1a or CNOT8/Caf1b. Other than being shown to have a role in deadenylation, knockdown experiments have shown that CNOT7 and CNOT8 also play a role in cell proliferation; interestingly, a double knockdown of CNOT7/8 has an increased effect on cell proliferation indicating redundancy in their function. Similarly, CNOT6/6L have also been shown to play a role in cell proliferation, but additionally, their knockdown also indicated a role in cell viability⁷⁷. Caf1 has been shown to interact with the antiproliferative proteins TOB1 and BTG2. In relation to deadenylation, BTG2 has been shown to enhance the efficiency of Caf1, however a Ccr4-Caf1-BTG2 trimeric complex is more efficient at deadenylation than Ccr4 or Caf1 alone^{74,75}.

The PAN2-PAN3 complex is formed of the PABP-dependent poly(A) nucleases PAN2 and PAN3 which as the name suggests, is a mechanism of deadenylation that requires the presence of polyA binding protein. PAN2 is a Mg^{2+} 3' to 5' exonuclease but has no ability to bind RNA. Instead, PAN2 forms a complex with a PAN3 homodimer via a linker region. PAN3 binds to RNA via the CTD and N-terminal zinc finger and is also able to interact with PABP and this interaction stimulates the RNase activity of PAN2^{76,77}. Once the PAN2-PAN3 complex shortens the tail to a length of 10-25 adenosine, the Ccr4-NOT complex removes the remaining tail and the remaining RNA is degraded by 3' to 5' decay by the exosome, or by decapping and 5' to 3' decay⁸².

1.2.5.2 Decapping

The removal of the 7-methylguanosine cap is an alternative mechanism for mRNA decay, whereby the cap structure is removed allowing for 5' to 3' exonucleases to act on the transcript. Deadenylation leads to the destabilisation of eIF4E from the cap structure and allows the Lsm1p-7p complex to interact with the m7G cap and recruit the decapping complex^{68,79}. The decapping complex consists of two proteins: Dcp1 and Dcp2. Dcp2 carries out the decapping function and is stimulated by Dcp1 as well as bridging factors such as Ge-1 and Hedls^{80,81}.

1.3: Types of RNA

For many years, RNA was merely thought of as the intermediate between DNA and protein synthesis, perhaps in some part due to the popularity of Crick's 'Central Dogma' theory. In 1961, mRNA was identified and, as it was shown to bind to ribosomes, was referred to as an 'intermediate carrying information from genes to ribosomes for protein synthesis'⁸⁶. However, in the past two decades, the more diverse roles of RNA have become more apparent, and RNA is no longer just the intermediate stage between DNA and protein.

The vast majority of transcription is not transcribing genes for protein production, but for transcribing non-coding (nc)RNA. Already mentioned are the small RNAs involved in translation, which are both transcribed by RNA polymerase III: tRNA and rRNA. Transfer RNAs (tRNAs) are between 76 to 93 nt in length and characterised by a highly conserved clover leaf secondary structure as well as a tertiary L shape structure which allows the

tRNA to interact with the ribosome⁵⁴. Ribosomal RNAs (rRNAs) are transcribed from large transcriptional units, around 43 kb in length, with individual rRNAs separated by an intergenic spacer⁸⁷.

Some ncRNAs are involved in post-transcription modifications of other RNA molecules, or DNA replication. TERC (telomerase RNA component) is involved in telomere replication. Small nucleolar (sno)RNAs are involved with modifying rRNA, such as methylation and pseudouridylation. Small Cajal body-specific (sca)RNA are snoRNAs which localise to Cajal bodies and are involved in the biogenesis of snRNPs (small nuclear ribonucleoproteins).

snRNPs consist of Sm proteins and small nuclear (sn)RNA and form the spliceosome, which is responsible for the removal of non-coding RNA called introns from mRNA. The snRNAs are commonly referred to as U-RNA as they have a high uridine content. How the spliceosome functions was outlined in Chapter **1.2.3.2**.

Another type of RNA is long non-coding (lnc)RNA. The main definition of a lncRNA is an RNA transcript that does not code for a protein and is longer than 200 nucleotides. lncRNA may be spliced, polyadenylated and capped, and have many of the same hallmarks found on mRNA⁸⁸ but recent work has also show that the transcription of lncRNA does differ to the transcription of protein-coding genes⁸⁹. Initially these non-coding transcripts were thought to have been generated by spurious transcription and not have a specific function^{90,91}. One of the first lncRNA to be identified as having a function is Xist (X_i-specific transcript) which is expressed from inactivated X chromosomes (X_i) rather than activated X chromosomes (X_a)⁹². Xist expression leads to transcriptional silencing and conformational reorganization of the X chromosome it is being expressed from⁹³.

Recently discovered are circular (circ)RNA. CircRNAs are RNA molecules that can act as a sponge for miRNA and RNA binding proteins, as well as playing a role in regulating splicing and transcription. The biogenesis of circRNA is still little understood, however the evidence suggests that they are formed by back-splicing, and there are currently four models describing how this could occur^{91,92}.

1.4: siRNA and miRNA

There are a variety of regulatory RNAs. Small interfering (si)RNA and micro (mi)RNA are both small molecules, typically 21-24 nucleotides in length which play a role in gene regulation. siRNAs are commonly used in cell culture to knockdown the expression of a gene of interest. siRNA and miRNA are identical in chemical structure, but have distinct differences in how they are produced and function. miRNAs are produced only from long endogenous transcripts transcribed in the nucleus and are initially processed by the Microprocessor before being exported to the cytoplasm as a double stranded duplex. siRNAs were initially thought to be produced purely by exogenous double-stranded RNA transcripts which have made their way into the cytoplasm, such as viral RNA. Repeated endogenous sequences, such as those found in heterochromatin have also been shown to be used in siRNA production⁸⁶.

While siRNA and miRNA originate differently, once in the cytoplasm they are both processed by Dicer and loaded on to RISC (RNA induced silencing complex) where one strand of the duplex is degraded while the other strand, called the guide strand, is used for targeting complementary RNA. siRNA binds to perfectly complementary RNA targets, and leads to endonucleolytic cleavage, which is followed by degradation of the target RNA. In mammals, miRISC does not require complete homology to function, and uses a small 'seed region' to target RNA, usually via 3'UTR sites. Once bound, miRNA represses translation and leads to mRNA degradation, however, perfectly complementary miRNA has been shown to lead to mRNA cleavage, as seen in siRNA^{91,92}.

1.5: Genomic locations of miRNA

According to the latest miRBase release⁹⁸ (release number 21, June 2014) over 2500 human miRNA sequences have been identified. Deep sequencing of total small RNA has led to this jump in novel miRNAs, but the majority of these newer sequences have little proof to show their function^{94,95}.

Classification of miRNAs by genomic location was first undertaken in 2004¹⁰¹. 161 miRNAs were analysed; 117 were found to be intronic, while 90 were located within the introns of protein coding genes and 27 were of long non-coding RNA (lncRNA) origin. 30 miRNAs had an exonic location, with the remaining 14 having a 'mixed' location, being affected by

alternative splicing of the host mRNA. These 161 miRNA were located within 123 host genes or transcriptional units (TUs), indicating that miRNAs may be clustered, with more than one miRNA being located within a single TU. This would mean that expression of miRNAs may be more complex than initially thought and led to a variety of new questions. 1) Are miRNAs in intronic and exonic loci processed in the same way? 2) Are introns spliced before the Microprocessor can function? 3) How are exonic miRNAs excised from miRNA and what effect does this have on protein coding transcripts? 4) Does miRNA expression from intronic and exonic loci differ?

A 2007 paper by the Kim group shows that 81 % of miRNAs have an intronic location¹⁰², which indicates that up to 20 % of miRNAs may be exonic. However, a 2015 paper used a trans-dominant negative (TN) Drosha, which is a Drosha protein mutated in the catalytic RNase domains, to prevent Microprocessing leading to the degradation of pri-miRNA transcripts; this would allow for better representation of pri-miRNA transcripts processed by Drosha in RNA-seq analysis. This data suggests that the majority of hairpins located within protein coding genes are located within introns (75 % in humans and 83 % in mouse). Interestingly, this work also looked at hairpins within lncRNAs and found that 39 % of human and 40 % of mouse hairpins were located in introns of lncRNA, whereas it was initially thought that miRNAs were mainly located in protein coding genes¹⁰³. This disparity in the amount of hairpins located within an intron in protein coding transcripts compared to lncRNA transcripts indicates that location of a hairpin may have an effect on the host transcript, as far fewer hairpins are found in coding exons compared to exons in lncRNA.

1.6: MiRNA

The discovery of miRNAs (miRNA) over two decades ago has led to major developments in understanding of the mechanisms behind gene regulation, and thus, led to the greater understanding of many biological processes fundamental to life. The first miRNA to be discovered was lin-4 in 1993¹⁰⁴. Lin-4 was found to regulate developmental processes in *Caenorhabditis elegans* by interacting with the 3' UTR of *lin-14* mRNA. In 2000, the second miRNA was discovered in *C.elegans*: Let-7¹⁰⁵. Like lin-4, let-7 also was found to play a role in regulating developmental timing in *C.elegans* and

also interacted with the 3' UTR of mRNA. It took 8 years from the discovery of lin-4 for the term 'micro-RNA' to be coined¹⁰⁶.

Let-7 and lin-4 were initially referred to as stRNA, or short temporal RNA. As they were the only stRNAs identified, and were only identified in *C. elegans* they were assumed to be an oddity for the species. This all changed when three papers were published in Science in 2001 which between them not only identified multiple miRNAs, but also found that they were highly conserved between species¹⁰⁷⁻¹⁰⁹.

The first papers defining miRNAs discovered the role of these short, ~22 nucleotide sequences during embryogenesis and organogenesis^{103,107} and since then a plethora of miRNA have been associated with an increasing number of diseases¹¹¹. With the importance of miRNA in disease now elucidated, it is vital to understand the role these miRNAs play and how they can be used to create efficacious drugs and improve disease diagnosis.

Currently, 35828 mature miRNA products have been identified in 223 different species with over 2500 miRNAs estimated to be expressed in human cells alone. This is large increase from the 232 mammalian miRNA identified in 2004¹⁰¹.

1.6.1: MiRNA Biogenesis

The biogenesis of miRNAs is a complex and highly regulated mechanism, processing a long primary RNA transcript (pri-miRNA) into mature miRNA via a pre-miRNA intermediate containing the characteristic stem-loop structure (**Figure 1.17**). Cellular localisation of pre-miRNA and mature miRNA indicated that the processing occurs in two steps; the first occurring in the nucleus by the Microprocessor complex, which consists of the proteins Drosha and DGCR8, and the latter in the cytoplasm¹¹². Two RNase III proteins have been shown to play a major role in miRNA biogenesis- the class II RNase protein Drosha and the class III Dicer.

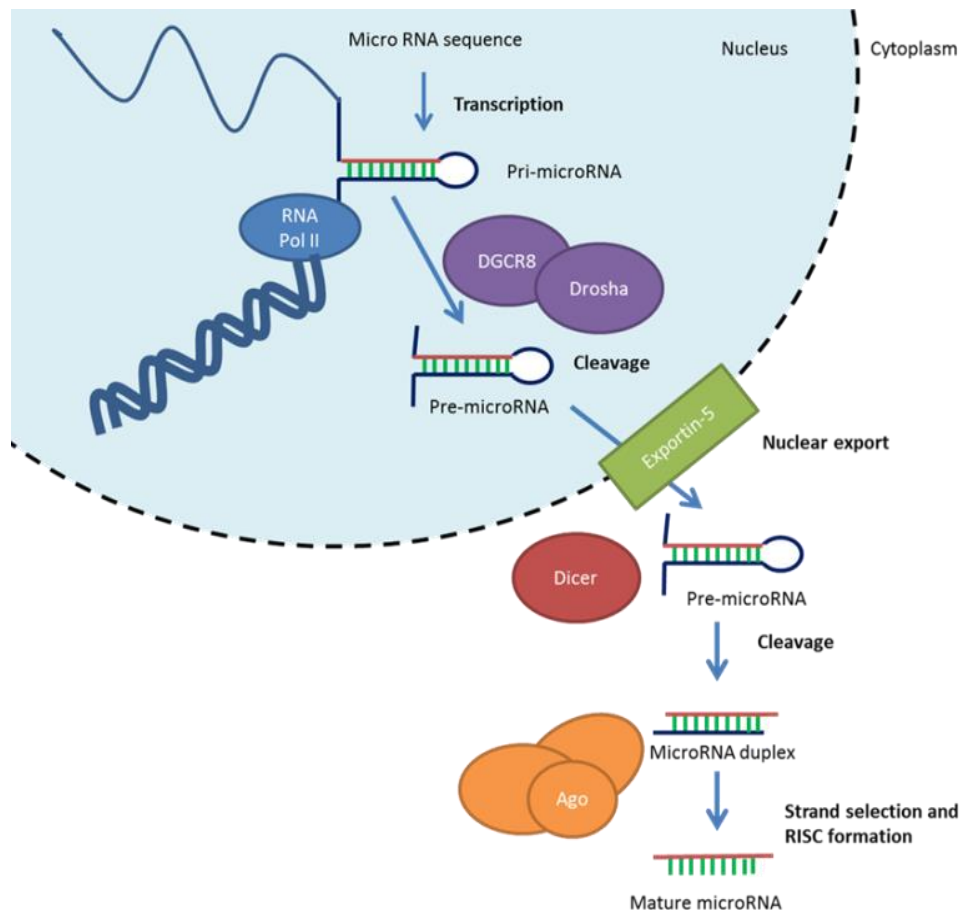


Figure 1.17: A schematic of miRNA biogenesis. The figure shows the production of mature miRNA from the point of transcription.

1.6.2: Transcription of miRNA

Initially, RNA polymerase III was thought to be involved in the biogenesis of miRNA as it is known to transcribe other small RNAs, such as U6 snRNA (small nucleolar RNA). However, work undertaken by Cai *et al* suggested that miRNAs are transcribed by RNA polymerase II (Pol II): pri-miRNA transcripts are long transcripts which have a similar structure to mRNA, and are capped and polyadenylated¹¹³. This work was also supported by another paper which also showed that Pol II interacts with the miRNA host gene promoters¹¹².

As mentioned in **1.2.3**, the carboxy-terminal domain (CTD) of RNA polymerase II provides a binding site for proteins involved in mRNA processing, such as splicing and 3' processing. Recent work has since indicated that the Microprocessor proteins interact with Pol II at the CTD¹¹⁴, adding to the complexity of complexes at the CTD and supporting the idea that links between miRNA processing and transcription occur.

1.6.3: Microprocessor Complex

The newly transcribed RNA containing a miRNA hairpin is subsequently processed by what is now known as the Microprocessor complex to release a pre-miRNA of approximately 70 nt which has a 2 nt 3' overhang. The first protein of the Microprocessor complex to have been identified is Drosha which is known to contain dsRNA binding domain and two RNase III domains (**Figure 1.18**)^{99,110}.

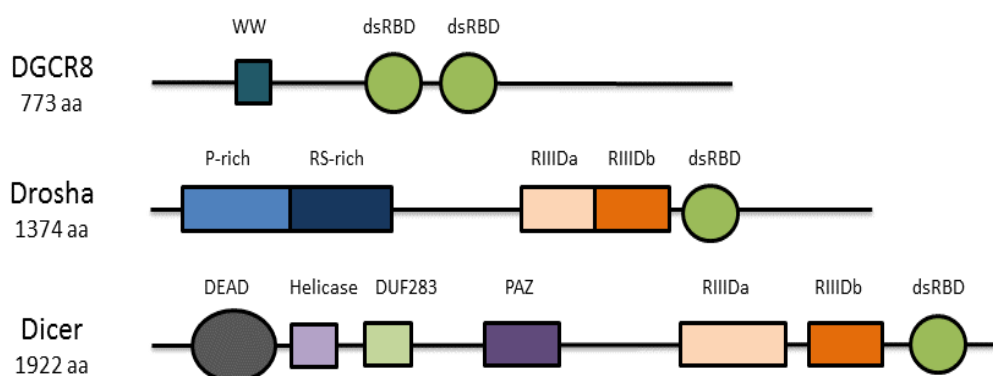


Figure 1.18: A comparison of Drosha, DGCR8 and Dicer domains.

Lee *et al* were able to show that Drosha is responsible for processing the pri-miRNA to the stem-loop miRNA precursor (pre-miRNA) by expressing and purifying FLAG-tagged Drosha and assessing its activity on in vitro transcribed radiolabelled pri-miRNA by gel analysis. This technique became ubiquitous among the research published at this early stage. By using FLAG-Drosha to pull out interacting proteins, it was shown that Drosha acts in complex with a second protein, DiGeorge Critical Region 8 (DGCR8)^{110,111}. Mutagenesis of the C-terminus of DGCR8 showed that this was the region that interacts with Drosha¹¹⁷.

1.6.3.1: Hairpin Recognition by the Microprocessor

How the Microprocessor interacts with pri-miRNA is still not completely understood, although some sequence elements of pri-miRNA that promote processing have been recently distinguished; it has been shown that miRNA processing is species specific and that a miRNA expressed in an unrelated species is not efficiently processed by the Microprocessor proteins¹¹⁸. From this, they showed that an 11 bp stem loop is optimal for effective processing and identified three primary sequence motifs that promotes processing in human pri-miRNAs; a CNNC domain downstream of the 5' Drosha cleavage site, a basal UG domain upstream of the 3' Drosha cleavage site and an UGUG domain in the apical stem and loop structure

(**Figure 1.19**). In addition, this work also showed that the CNC domain has a role in splicing efficiency by interacting with SRp20, a known protein in splicing recognition. METTL3 (methyltransferase-like 3) has also been shown to play a role in the recognition of hairpins by the Microprocessor: METTL3 methylates adenosine nucleotides, producing N⁶-methyladenosine or m⁶A, and targets pri-miRNA regions which may contribute to recognition by DGCR8¹¹⁹.

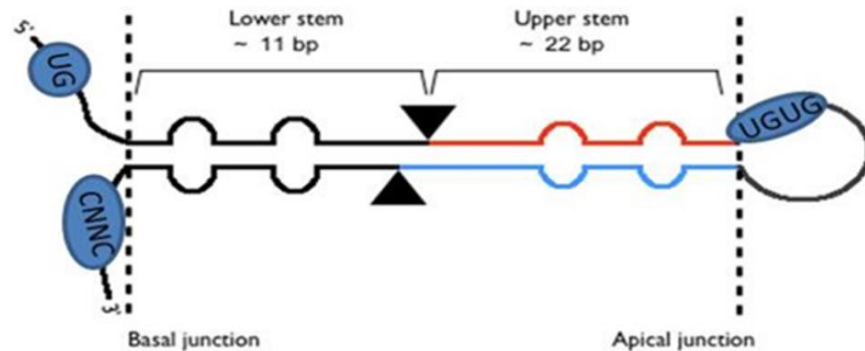


Figure 1.19: A structural representation of a miRNA hairpin. A diagram depicting the secondary structure of a miRNA hairpin. Blue circles show known common sequence motifs. Black arrows represent the Drosha cleavage sites.

Structural analyses of Drosha and DGCR8 have also provided some insights. As a class II RNase III super-family protein, Drosha contains two RNase III binding domains that are thought to create an intramolecular dimer which can process pri-miRNA by acting as a catalytic site¹²⁰. The importance of the terminal loop has also been detailed, with work showing that a loop of at least 10 bp is needed for efficient processing¹²¹. The same paper also underlined how the length of the stem can affect the position of Drosha cleavage in relation to the terminal loop, showing that the addition or removal of a single base shifts the Drosha cleavage site by the same amount. This suggests that Drosha, on recognising the stem sequence, cleaves at a set distance, rather than recognising a specific sequence. The theory that Drosha acts as a molecular ruler is further confirmed by a paper by Nguyen *et al* which shows that Drosha interacts with the basal ssRNA-dsRNA junction of a pri-miRNA stem loop, specifically the UG motif¹²².

A later paper used a combination of stem and terminal loop mutations and found that pri-miRNA with a shorter stem was still cleaved at the original

site, with a reduction in efficiency¹²³. While this may disagree with the work by Zhang¹²⁰, both papers show that the stem is important for cleavage, while the size of the terminal loop can affect processing efficiency. Han *et al* also showed that it is DGCR8, not Drosha, which interacts with the pri-miRNA¹²³. A second paper then showed that it is the dsRDB subunits that provide this binding ability¹¹⁷. By comparing the binding efficiency of both pri- and pre-miRNA to DGCR8, they have been able to show that DGCR8 is able to recognise and bind to sequences outside of the stem-loop structure found in the ssRNA section of the pri-miRNA. This work led to the ssRNA-dsRNA junction anchoring model, underpinning how the Microprocessor proteins work together, with DGCR8 being involved in substrate recognition, and aiding Drosha cleavage at the correct location.

With the Microprocessor proteins elucidated, understanding the relationship between Drosha and DGCR8 was then paramount. Knockdown analysis of the two proteins by siRNA (small interfering RNAs) seemed to have different effects; Drosha depletion led to an increase in the levels of DGCR8, while DGCR8 depletion also leads to a decrease in Drosha levels as DGCR8 has been shown to stabilise Drosha at the protein level^{116,124,125}. It has also been shown that the DGCR8 mRNA contains a hairpin in the 5' UTR that is processed by the Microprocessor, leading to a post-transcriptional feed-back loop¹²⁵.

1.6.3.2: Modification of DGCR8

23 phosphorylation sites have been found in the N-terminus of DGCR8 and have been shown to increase the stability of the protein. However, DGCR8 phosphorylation was shown to have no effect on Drosha/DGCR8 interactions or Microprocessor activity¹²⁶. The N-terminus is important for proper nuclear localisation¹¹⁷ although it has been shown that phosphorylation may not affect these processes¹²⁶.

As well as being phosphorylated, DGCR8 can also be SUMOylated at K707 by SUMO1 protecting DGCR8 from degradation by preventing ubiquitination. While this work shows that SUMOylation does not affect Microprocessor activity, it does increase DGCR8 affinity to pri-miRNA¹²⁷. Lysine 259 can also be SUMOylated, and promotes the nuclear localisation

of DGCR8¹²⁸. Both of these SUMOylation sites have been shown to play a role in tumour suppression^{127,128}.

1.6.3.3: Microprocessor stoichiometry

To understand how the Microprocessor functions, it is necessary to understand the relationship between the two Microprocessor proteins: Drosha and DGCR8. Using purified monomers and dimers of DGCR8, it has been shown that dimerised DGCR8 leads to more efficient processing of pri-miRNA¹²². Similarly, photobleaching assays have also indicated that the Microprocessor consists of two DGCR8 molecules to one Drosha, but shows a 1:1 ratio is also possible¹²⁹. It has been shown that it is the Rhed (RNA-binding heme domain) of DGCR8 that specifically binds to pri-miRNA at the apical and basal sites of the hairpin. DGCR8 binding to the UGU motif in the apical region of the hairpin has been shown elsewhere¹²². Each DGCR8 protein contains one Rhed domain, indicating that two Rhed domains, and therefore two DGCR8 proteins, are required for microprocessing¹³⁰ and other work has shown that the Rhed domain requires Fe³⁺ to function efficiently¹³⁰⁻¹³².

1.6.3.4: Co-transcriptional Microprocessing

Microprocessing has been shown to occur co-transcriptionally. Using chromatin immunoprecipitation (ChIP), it was shown that Drosha associates with chromatin-associated miRNA hairpins, and that DGCR8 may interact with the pri-miRNA hairpin prior to Drosha. The co-transcriptional function of Drosha processing was further shown using NRO (nuclear run-on) experiments using HeLa cells transfected with miR-330 into intron 2 of a B-globin gene. This experiment showed that Drosha cleavage occurs co-transcriptionally, and that Drosha cleavage is a relatively fast process which occurs before splicing. In addition, this work also showed that introns processed by Drosha are co-transcriptionally degraded by XRN2¹³³.

Work by the Kim lab has also shown that Drosha can process miRNA hairpins from introns of mRNA prior to splicing¹⁰². This work showed that splicing of an intron is not needed for the cleavage of a miRNA from an intron, and that Drosha can even process partially spliced substrates. Using Drosha knockdown experiments, it was also shown that the Microprocessor may interfere with splicing. Not only did Drosha knockdown lead to an

increase in the levels of spliced mRNA when a miRNA was present in an intron, but introns adjacent to an intron containing a miRNA may be spliced more rapidly.

The effect of Drosha cleavage on downstream transcription has also been examined¹³⁴. A construct containing two miRNAs, miR-223 and miR-196a-1, within a single transcript was transfected into HeLa cells. The abundance of the miRNAs were measured by qPCR, which showed that the processing of miR-223 affected the downstream processing of miR-196a-1. Increasing the distance between the two miRNAs further decreased the levels of miR-196a-1 in relation to miR-223. ChIP analysis using RNA polymerase II antibodies showed that Drosha cleavage can lead to transcriptional attenuation, as less RNA polymerase was detected downstream of the first Drosha cleavage site.

The binding of both Microprocessor proteins to RNA has been shown to be transcription dependent, with Drosha and DGCR8 shown to interact with the 5' end, or promoter-proximal regions, of protein coding genes, and that this occurs independently of miRNA biogenesis or RNA cleavage¹¹⁴. Using knockdown experiments, Drosha was shown to have a positive effect on transcription, with Drosha depletion leading to a decrease in the levels of pre-mRNA. This work also showed that Drosha interacts directly with RNA polymerase II and CBP80 via its N-terminal domain. This adds further evidence to the co-transcriptional effects of the Microprocessor.

1.6.3.6: Microprocessor efficiency

Recent published work has shown that some miRNAs are processed faster than others. This is true for polycistronic miRNAs which are a cluster of miRNAs which are part of the same transcriptional unit. Using a pulse-chase experiment looking at newly transcribed RNA, some miRNA were shown to be processed within the pulse, whereas the majority of intermediate miRNA were processed within 15-30 minutes and completely processed by 60 min. Interestingly, the polycistronic miR-221/222 transcript which contains two miRNAs was shown to process miR-222 faster than miR-221. However, the mature levels and the half-life for both these miRNA are similar. Late processing miRNAs were only 50 % processed by 60 mins. The data indicates that slow and intermediate miRNAs are less likely to be processed co-transcriptionally¹³⁵. While this

work does not look at where slow or intermediate miRNA may be being processed, it is possible that microprocessing may be occurring in the nucleoplasm once released from the chromatin.

1.6.3.7: Drosha localisation

It is also possible that pri-miRNA may be processed in the cytoplasm. Work by Dai *et al* has shown that Drosha can be found in both the cytoplasm and nucleus and that this may be due to alternative splicing. Exon 6 of Drosha contains a nuclear localisation signal (NLS) and exon 7 contains a nuclear exportation signal (NES) with Drosha isoforms lacking one of, or both, exon 6 and 7, while the full length Drosha can also be detected in human tissue. This work also showed that pri-miRNA are present in the cytoplasm¹³⁶. Similar work has also shown the importance of alternative splicing of Drosha pre-mRNA when it comes to the localisation of Drosha and indicates that there are two regions where alternative splicing may occur. The first is in the 5' UTR, with some transcripts lacking exon 2, and some transcripts having a shorter first exon due to a second TSS. As indicated by the work by Dai *et al*, the second region of alternative splicing is alternative inclusion of exon 6 or exon 7. This work showed that all the varieties of Drosha isoforms can be expressed in multiple human tissues, can bind to DGCR8, and process pri-miRNA. Interestingly, it confirms that isoforms containing exon 7 could localise to both the nucleus and cytoplasm, indicating that the NES present in exon 7 only weakly locates to the cytoplasm.

1.6.4: Nuclear Export

Exportin-5 (Exp5), a member of the karyopherin family of nucleocytoplasmic exporter with Exp5 knockdown leading to a significant reduction of mature miRNA in the cytoplasm. This indicates that Exp5 is responsible for pre-miRNA export out of the nucleus¹³⁷.

However, recent work has shown that Exp5 may not be as important when it comes to the nuclear export of pre-miRNA as initially thought. A knockout experiment shows a small effect on miRNA expression, although it may have an effect on the degradation on the 3' end of the 3p miRNA. This indicates that other Exportins may also be able to export pri-miRNA into the cytoplasm, and that Exp5 may only be important for a subset of miRNAs, such as let-7¹³⁸.

1.6.5: Dicer

Pre-miRNA exported into the cytoplasm is then subjected to further processing, cropping the ~70 nucleotide strand into a ~21 nucleotide duplex which is unwound to release the mature single-stranded miRNA. Dicer, an RNase III protein suspected of carrying out this function, and showed that it had some interaction with the RNAi-induced silencing complex (RISC) which at this time was known to be the effector of this silencing process¹³⁹. Dicer has been shown to contain a dsRNA (double-stranded RNA) binding domain and two RNase III domains as well as a helicase and PAZ domain either side of an N-terminal domain (**Figure 1.18**).

The 3' overhang produced by Microprocessing is important for Dicer processing with work showing that Dicer requires a free end, rather than a closed structure such as a loop, and that a length of two nucleotides results in efficient processing¹⁴⁰. This work also showed that the 3' end of the pre-miRNA is used as a 'ruler' ultimately determining the length of the mature miRNA, now referred to as the 3' counting rule.

1.6.7: RNA-induced silencing complex (RISC) function

The RISC consists of Dicer, an Ago protein, and a dsRBP such as TRBP^{139,140,143} in a 1:1:1 ratio¹⁴⁴. While Dicer is known to process the pre-miRNA into a miRNA duplex only one of the two RNA strands, namely the guide strand is used in silencing, while the passenger strand, which is often annotated as miR*, is degraded¹⁴⁵. The use of deep sequencing to analyse global populations of miRNA have shown that many passenger strands are functional, and in some cases the passenger strand may be more highly expressed than the guide strand¹⁴⁶. This has led to a change to the nomenclature with 5p and 3p used to describe which arm the mature miRNA has come from. This slicing activity is mediated by Ago proteins. In humans, there are four Ago proteins which are known to associate in RISC, with only Ago2 able to cleave targeted mRNA¹⁴⁷.

TRBP was also shown to recruit Ago2 to RISC and stabilise Dicer in complex¹⁴⁸ and that Dicer 'gives' the guide strand to Ago2¹⁴⁴. More recently, PACT was shown to bind dsRNA and form complexes with Dicer and Ago2, just like TRBP, but was shown to have an inhibitory effect on RISC¹⁴⁹. The large number of proteins shown to be able to interact with the

constituents of RISC just goes to indicate the fluidity of the complex and hints at how the complex itself could be regulated.

1.6.7.1: RISC loading

The 5' ends of the miRNA duplex is important in RISC loading¹⁵⁰. The 5' end which forms the least stable base pairing to the other side of the duplex is able to be unwound by RISC and efficiently loaded while a stable 5' end prevents unwinding and the 5' antisense strand has been shown to be generally less thermodynamically stable. From this work it was suggested that the ratio between guide and passenger strands loaded onto RISC should be constant although this ratio could be affected by modifications to the 5' end and variations of stability across different miRNAs. This work was confirmed by another group who also showed Ago3 preferentially interacts with the passenger strand indicating that the passenger strand has a yet undiscovered function¹⁵¹. While this work focuses more towards siRNA, the same rules apply with respect to miRNA, but both strands are incorporated into RISC to some extent.

1.6.7.2: Target recognition

The guide miRNA has been shown to act more like an RNA binding protein than free RNA, using the seed region to probe for target sequences¹⁵². Target recognition by the miRNA seed region has also been shown to cause conformational changes in Ago2 structure. Firstly, 2-5 nt of the seed region bind, leading to changes in Ago2 structure, allowing nucleotides 2-8 and 13-16 of the miRNA to be exposed to ensure specificity¹⁵³. This was also shown by a different group, which also showed that RISC recognises target sequences using the 2-8 nt seed region, much like probe, with the 5' region of the target miRNA matching first before extending to the 3' to form a stable association¹⁵⁴. This work also showed that RISC is able to bind to complementary DNA but rapidly dissociates.

While conformational changes have been shown to occur in miRNA-associated Ago2, the conformation of Ago2 can affect how miRNA sits in the protein. Ago2 has a range of partially open states, as well as an open and closed state. miRNA was shown to bind to Ago2 differently depending on the conformational state¹⁵⁵.

1.6.7.3: RISC-mediated gene silencing

There are two mechanisms by which miRNA-induced silencing complex function to regulate gene expression: translational repression and mRNA decay. The current evidence suggests that both mechanisms do occur, but translational repression is likely to occur first.

RISC is dependent on the association of GW182 to Ago2. In total, humans have three homologs of GW182: TNRC6B, TNRC6C and GW182 (TNRC6A). The GW182 functions are redundant with respect to RISC activity, although there is some evidence that GW182 and TNRC6B have a role in miRNA stability, with knockdown experiments leading to a reduction of miRNA being secreted via exosomes, but not for TNRC6C¹⁵⁶. GW182 has also been shown to protect RISC loaded-miRNA from degradation from nucleases^{153,154}.

There have been multiple reports showing that GW182 interacts with the deadenylase complex CCR4-NOT¹⁵⁸⁻¹⁶¹ and the GW motif has been shown to contain two Ago binding sites¹⁶². GW182 has also been shown to recruit the PAN2-PAN3 deadenylation complex in a PABP-dependent manner¹⁶³.

As well as interacting with the CCR4-NOT and PAN2-PAN3 deadenylation complexes, GW182 has also been shown to interact with decapping factors. A transcript containing a miRNA binding site was engineered to be transcribed by RNA polymerase III and therefore would not have a polyA tail or a cap; in *Drosophila*, decapping factors were still recruited to the transcript showing that decapping factors are recruited to RISC independent of a cap or a polyA tail¹⁶⁴.

While the evidence is clear that RISC leads to mRNA decay by deadenylation and decapping, RISC has also been shown to lead to translational repression, with translational repression being shown to occur before, and is a prerequisite for deadenylation^{158,165}. However, how RISC causes translational repression is still unclear¹⁶⁶.

1.7: Regulation of miRNA biogenesis

The production of miRNA is highly complex and regulated by a range of factors from transcription to RISC loading. Some, but not all of the methods of regulation will be discussed below.

1.7.1: M6A methylation

Adenosines can be post-transcriptionally modified by METTL3 which adds a methyl group to the N6 position of the adenosine molecule (m6A). The m6A modification is found throughout the genome, and has been found in over 7600 genes and over 300 non-coding RNAs¹⁶⁷.

METTL3 recognises an RGAC motif; this motif is highly enriched in pri-miRNAs, but not pre-miRNA. METTL3 depletion was shown to lead to a global decrease in the levels of mature miRNAs as well as an increase in the levels of unprocessed pri-miRNA¹¹⁹. Subsequent experiments showed that DGCR8 interacts with methylated RNA, with METTL3 knockdown leading to a reduction the amount of RNA bound to DGCR8. This work shows that m6A methylation aids Microprocessor recognition. In addition, the genomic locations of the miRNAs downregulated following METTL3 depletion were also analysed, with intronic miRNAs predominately affected. This is not surprising considering the majority of miRNA are located within introns.

1.7.2: Cofactors

As mentioned in chapter **1.6.3.5**, Drosha can interact with the cap binding complex (CBC) protein CBP80¹¹⁴. The CBC has also been shown to have a role in the processing of pri-miRNAs, with *cbp80* and *cbp20* mutants in Arabidopsis shown to have higher levels of pri-miRNAs. These mutants also a decrease in the levels of spliced mRNAs, with the first intron less efficiently spliced in particular. This indicates a role of the CDC in both miRNA biogenesis and splicing¹⁶⁸.

A component of the CBC, Ars2 has also been shown to have a role in miRNA biogenesis. HeLa cells were treated with siRNA to target Ars2 or CBP80 and qPCR analysis showed that the levels of mature let-7 and miR-21 were reduced, although not as significantly reduced as compared to DGCR8 knockdown. Co-immunoprecipitation showed that Ars2 interacts

with Drosha leading to the idea that Ars2 and the CBC complex may be involved in pri-miRNA processing¹⁶⁹.

In *Drosophila*, immunoprecipitation showed that Ars2 has also interacts with Dcr-2 in the siRNA biogenesis pathway, although no interaction was seen for Dcr-1. In the miRNA biogenesis pathway, this work showed that Ars2 interacts directly with the Microprocessor leading to the stabilisation of the pri-miRNA¹⁷⁰.

The RNA binding protein FUS (also called FUS/TLS- fused in sarcoma/translocated in liposarcoma) has been shown to be important for the expression of a subset of miRNAs by binding to pri-miRNA. This work also showed that FUS interaction is necessary for the efficient recruitment of Drosha for some pri-miRNAs¹⁷¹.

The kinase GSK3 β has been shown to phosphorylate Drosha at two points on its N-terminus: serine 300 and serine 302. Phosphorylation at these serines leads to Drosha being translocated to the nucleus^{172,173}. Inhibition of GSK β has been shown to reduce the activity of Drosha and ultimately lead to a reduction of pri-miRNA processing and mature miRNA expression. Further experiments showed that GSK β may aid the interaction between Drosha and DGCR8 and help facilitate their binding. This work also showed that GSK β also interact with another known Microprocessor cofactor p72¹⁷⁴.

DDX17 is a DEAD-box helicase and is also referred to as p72 in mice. Crosslinking experiments have shown that p72 is able to bind pri-miRNA hairpins¹⁷⁵. Other work has shown that p68 (DDX5 in humans) and p72 fractionate with the Microprocessor complex, and are required to process a subset of pri-miRNA¹⁷⁶. Both DDX5 and DDX17 have been shown to have a role in alternative splicing which could indicate a mechanism for how miRNA biogenesis and splicing interact¹⁷⁷. In addition, P72, and KHSRP (KH-type splicing regulatory protein, an RNA binding protein), but not p68, have been shown to stabilise the protein level of unloaded Ago2¹⁷⁸.

The tumour suppressor gene P53 is mutated in 50 % of cancers, and mutant P53 has been shown to modulate the biogenesis of a subset of miRNAs. Mutant P53 has been shown to bind and sequester p72 and p82,

which will affect the association of Drosha to pri-miRNAs and leads to an inhibition of mature miRNA expression¹⁷⁹.

The Hippo signalling pathway has also been shown to regulation miRNA biogenesis. When cells are at confluency, YAP is retained in the cytoplasm with allows p72 to associate with the Microprocessor. When the cells are less confluent, YAP is found in the nucleus and sequesters p72. Therefore, cell confluency can play a role in the regulation of miRNA expression¹⁸⁰.

1.8: miRNA turnover

When incorporated into RISC, mature miRNAs have been shown to be very stable^{181,182}. Studies have shown that the degradation of miRNA can be modulated in a physiological relevant way and can be another method of regulating miRNA expression¹⁸³.

Work in plants showed that HEN1 interacts with miRNA duplexes, and recognises and methylates the terminal nucleotide of the 3' overhang^{184,185}. The methylation of miRNAs prevents uridylation, with *HEN1* mutants leading to an increase in the uridylation of miRNAs¹⁸⁶. Uridylation is carried out by HESO1, which leads to the degradation of miRNA¹⁸⁷. In humans, terminal uridylyl transferases (TUTs) such as TUT2, TUT4 and TUT7 are responsible for the uridylation of miRNAs. 3' adenylation mediated by GLD-2 has been shown to increase the stability of miR-122 in the liver¹⁸⁸.

The mature miRNA expression of the miR-16 family, and in particular miR-503, is highly dynamic and has been shown to respond to the cell cycle with these miRNAs having unusually short half lives. This instability has been shown to be mediated by the miRNA sequence: the seed region and the 3' end¹⁸⁹. The miR-29 family have also been shown to contain sequence motifs that affect their stability. MiR-29a is a stable miRNA, whereas, miR-29b has been showed to decay rapidly. These miRNAs only differ at two regions, one in the 3' region and the other at position 10. MiR-29b contains three uracils at position 9-11, and in miR-29a the central region is a cytosine. Interestingly, the uracil-rich region was shown to lead to rapid turnover of the miRNA¹⁹⁰.

There is also evidence which shows that miRNAs can be modified once they have bound to their target mRNA, which may lead to their degradation. In

Drosophila, a reduction in the levels of mature miR-277 was observed when the expression of its target mRNA was increased. To understand this effect further, radiolabelled let-7 duplex was added to *Drosophila* embryo lysate which contained let-7 target mRNA. The data indicated showed that the levels of let-7 decreases, however the levels of shorter (trimmed) or longer (tailed) let-7 increased over time¹⁹¹.

1.9: Splicing and miRNA biogenesis

The vast majority of miRNA sequences reside in non-coding locations in the genome, be this within an intron, or as a lncRNA¹⁹². Due to the better annotation of host genes, intronic miRNAs have been the focus of miRNA biogenesis research, although little is known about miRNA biogenesis from a lncRNA. To understand if exonic and intronic miRNA are processed differently, it is logical to first look at the expression of intronic miRNA via the Microprocessor.

Using Drosha siRNA to knockdown Drosha levels, small increases in spliced mRNA were identified¹⁰². This work also showed that introns containing miRNAs are spliced slower than introns lacking miRNA sequences, as well as that an unspliced intron can still be a Microprocessor substrate, indicating that the splicing machinery and Microprocessor may compete for substrate. This was further substantiated by a second group whose work showed that the Microprocessor cleaves before splicing has occurred¹³³. It was also shown that exonucleases associate with the Drosha cleavage site co-transcriptionally and promote intron degradation before splicing has occurred. A year later more evidence that the spliceosome and Microprocessor interact was published, showing that they cofractionate and that Drosha cleavage is kinetically faster than splicing¹⁹³.

With a possible relationship between the Microprocessor and the splicing machinery being shown^{128,198}, understanding how these two processing complexes interact became the next challenge. They were able to show that splicing factors recruit Drosha such as U1; U1 binds to the 5'SS before microprocessing begins and inhibiting splicing factor U1 reduces the global levels of intronic miRNAs, thus indicating that microprocessing and splicing of intronic miRNAs could be mechanistically coupled. Janas *et al* also showed that Drosha cleavage of intronic miRNAs promotes splicing of the host intron. This interaction was confirmed by another group who showed

that primary miRNA transcripts co-fractionate with the splicing machinery and that splice inhibition increases miRNA expression while the converse, Drosha knockdown, increases splicing¹⁹⁵.

A relationship between splicing and miRNA biogenesis has been identified in the nematode. Pri-let-7 was found to have two transcriptional start sites which produces nascent pri-let-7 as well as a trans-spliced isoform, which is where structural rearrangements of the transcript occur to promote Microprocessor cleavage and splicing. The trans-spliced isoform of pri-let-7 remodels the pri-let-7 secondary structure but it still looks like canonical pri- substrate. This work showed that the 3' splice site is necessary for efficient processing of the pri-miRNA and also showed that Pasha, the *C.elegans* homolog of DGCR8, binds preferentially to spliced pri-let-7. This work indicates that splicing may improve Microprocessor recognition¹⁹⁶.

The 11th exon of the protein-coding gene FSTL1, which is involved in wound healing, contains the miR-198 hairpin. Following epidermis wounding, both FSTL1 mRNA and its protein were expressed, while miR-198 was only detected prior to wounding. This suggests that miRNA expression from an exon is somehow regulated with respect to splicing. This regulation was found to be mediated by KHSRP, a KH-type splicing regulatory factor¹⁹⁷.

1.10: Non-canonical processing of miRNAs

Not all miRNAs are processed following the standard model of miRNA biogenesis. Drosha knockout mutants have been shown to have a severe reduction in global miRNA expression, however, there are some miRNAs that are unaffected by Drosha knockdown, and unexpectedly, some miRNAs that show an increased expression. Interestingly, Drosha and Xpo5 knockouts lead to an increase in the levels of Dicer. Dicer knockouts were shown to have less of an effect on global miRNA levels than Drosha knockouts which may be due to the ability of Ago to load pre-miRNA without Dicer cleavage¹³⁸. This work indicates that many miRNAs may be being processed non-canonically.

When it comes to miRNAs located across intron-exon junction (SO miRNA), it would be expected that the biogenesis of these transcripts would differ to

purely intronic miRNAs. By looking at the positions of miRNA within the primary sequence, which showed that while the 5' or 3' ends may overlap with splice sites, the terminal loop was shown not to overlap splice sites¹⁹⁸. Knockdown of the Microprocessor proteins showed an increase in splicing, while overexpression reduces levels of spliced mRNA. This pattern of splicing during knockdown and over-expression is similar to published work from the Sperling lab for mRNA in introns¹⁹⁵. However, SO miRNAs are located within an intron-exon junction which may lead to an additional regulatory step to mediate the interaction between miRNA biogenesis and splicing.

Small introns containing miRNAs have been discovered which have been shown to bypass Drosha processing. The miRNA, like normal intronic miRNAs consists of a hairpin, but the stem structure ends at the splice sites. When spliced, this would produce the characteristic 2 nt- 3' overhang, and be available for Dicer processing. The mature miRNA produced by this alternative mechanism still leads to silencing of the target mRNA¹⁹⁹. MiRNAs processed in this manner are called mirtrons.

A second mechanism of miRNA biogenesis which bypasses Drosha cleavage, but is still processed by Dicer, has also been identified. Where the 5' end of the pre-miRNA sequence overlaps with a transcriptional start site (TSS) the pre-miRNA gains a 5' cap, with the pre-miRNA being generated by transcription and not by Microprocessing. The presence of the 5' cap also causes the pre-miRNA to be exported not by Exp-5 but by Exp-1 similar to the export of capped mRNAs. The pre-miRNA is still processed by Dicer, but the 3p arm is preferentially loaded on to Ago2, while the capped 5p arm is targeted to degradation²⁰⁰.

A Dicer independent pathway has also been identified. Pre-miRNA-451 consists of a hairpin that when cleaved by the Microprocessor, creates a stem of only 17 nt, rather than the 19 nt stem required for dicer processing. In this instance, the hairpin is loaded directly on to Ago2, and the Ago2 catalytic site trims the 3' end²⁰¹.

1.11: Non-canonical functions of the Microprocessor

The functions of the Microprocessor proteins Drosha and DGCR8 are well explored in relation to miRNA biogenesis. However, there are reports of non-canonical functions of these proteins.

DGCR8 has been shown to bind to a variety of RNA types, not just miRNA hairpins. In one publication, the majority of DGCR8 binding was shown to bind to protein-coding mRNA and secondly to lncRNA which lack miRNAs such as MALAT1²⁰². However, a second paper looking at Drosha binding sites indicates that the majority of Drosha binding is on miRNA regions, but it is also found bound to other RNA sequences, particularly introns²⁰³. Both these papers show that the absence of DGCR8 or Drosha leads to an increase in non-miRNA sequences bound by the remaining Microprocessor protein indicating that the Microprocessor may have a non-canonical function that still leads to mRNA repression^{202,203}. One example of this would be non-miRNA exonic hairpins; the Microprocessor is able to interact with exonic hairpins in protein-coding genes, including DGCR8 mRNA, with DGCR8 knockdown leading to an increase in the mRNA. In this case, the Microprocessor is able to cleave the hairpin and the small RNAs produced are bound to Ago, however, so far they have been shown to have little or no biological relevance. The Microprocessor activity on exonic hairpins within protein-coding genes leads to the destabilisation of the transcript, which therefore leads to regulation of the gene²⁰³.

Knockout DGCR8 mESC cells created by CRISPR-Cas9 mutagenesis arrested at G1 phase and failed to express differentiation markers as well as preventing canonical miRNA expression. Interestingly, a phosphomutant DGCR8 recovered miRNA expression, but failed to rescue the cells from a pluripotent state. This confirms other published work that showed that the phosphorylation of DGCR8 is not required for miRNA biogenesis¹²⁶, but does indicate that the phosphorylation state of DGCR8 is required for a non-canonical function. Tcf7l1 contains an intronic hairpin which is recognised by wild type DGCR8 but not the phosphomutant; DGCR8 was shown to have a role in the alternative splicing of Tcf7l1, with the phosphomutant being shown to lead to aberrant isoforms²⁰⁴. This work clearly indicates a role of DGCR8, and possibly Microprocessor cleavage, on splicing efficiency.

DGCR8 has also been shown to interact with repetitive sequences, including transposons. Alu and LINE-1 sequences are negatively regulated by Microprocessor cleavage²⁰⁵.

1.12: miRNA in long non-coding RNA

As mentioned earlier in **1.5**, miRNAs can be found throughout the genome and many miRNAs are located within lncRNA¹⁰¹.

As discussed in chapter **1.2.3.1**, the cleavage and polyadenylation (CPA) mechanism is involved in the termination of transcription of RNA polymerase II transcripts²⁰⁶. Work focusing on pri-miR-122, a lncRNA which contains a hairpin in the terminal exon, shows that this mechanism is not used for a subset of lnc-pri-miRNAs. PolyA fractionation showed that pri-miR-122 transcripts were not polyadenylated even though the transcript does contain a PAS. Northern blotting showed that the spliced and unspliced transcripts were shorter than would be expected if cleavage occurred at the PAS, and indicated that transcription was being terminated upstream of this signal. Microprocessor knockdown using siRNA targeting DGCR8 led to an increase of transcriptional read through, indicating that transcriptional termination of pri-miR-122 is mediated by the Microprocessor. Deep sequencing showed that Microprocessor-mediated transcriptional termination is not unique to pri-miR-122, and that it occurs on a range of miRNA containing lncRNAs including pri-miR17~92a and pri-miR-29a~29b-1.

1.13.1: miR-122

The miRNA, miR-122, is highly liver specific, with 72 % of the miRNA pool in mouse livers shown to be miR-122²⁰⁷. Mature miR-122 is expressed from a lncRNA which is capped and not polyadenylated²⁰⁶. Pri-miR-122 consists of two exons, with the miRNA hairpin located at the end of the terminal exon (**Figure 1.20**).



Figure 1.20: A schematic of pri-miR-122.

MiR-122 has been implicated in a variety of diseases. MiR-122 has been shown to play a role in hepatocellular carcinoma, with a decrease in expression shown to correlate with a poor prognosis²⁰⁸. Interestingly, miR-122 can also be detected in the blood, and is used as a biomarker of liver injury, and non-alcoholic fatty-liver disease²⁰⁹. MiR-122 also has an important role in Hepatitis C replication.

Mature miR-122 is a highly liver specific miRNA which is expressed in the human hepatocyte derived cell line Huh7, but not HepG2 cell lines. Similarly, HCV can only replicate in Huh7 and not HepG2 cell lines. Work published by Sarnow group in 2005 looked at the role of mature miR-122 in HCV replication and found that the HCV genome contains two regions which would be able to Watson and Crick pair with the seed region of miR-122. One of these binding sites was a highly conserved region in the 3' non-coding region (NCR) whereas the second binding site was surprisingly found in the 5' NCR. Using 2'OMe to sequester the mature miR-122, the authors were able to show that miR-122 is necessary for HCV replication²¹⁰.

The expression of mature miR-122 is biologically important as miR-122 has been shown to lead to an increase in the levels of HCV virus replication^{211,212,217,218}. Inhibition of miR-122 using a linked nucleic acid (LNA)-antimiR-122 called SPC3649 which sequesters miR-122, has been shown to reduce the viral load of chimpanzees with a chronic HCV infection^{219,220}. SPC3649 has since been tested on humans, and has been renamed Miravirsen and has also been shown to reduce viral HCV levels²¹⁶.

1.13.2: miR-17~92

The human miR-17~92a cluster contains 6 miRNA sequences (**Figure 1.20**): miR-17, miR-18a, miR-19a, miR-20a, miR-19b-1, and miR-92a²¹⁷ and has two known orthologues, miR-106b~25 and miR-106a~363. Interestingly, the miR-17~92 cluster forms a globular 3D structure that prevents the Microprocessor accessing some of the hairpins. This leads to uneven expression of the six miRNAs with miR-17, 19a and 20 being more efficiently expressed than miR-18a, 19b and 92, with miR-17 shown to have an increased expression of 17 fold compared to miR-92a²¹⁸.



Figure 1.20: The order of the hairpins in pri-miR-17~92a.

The expression of the mature miRNAs from the cluster has also been shown to require spliceosome components for biogenesis to create an intermediate progenitor (pro)-miRNA. The pro-miRNA is approximately 800 bp in length, and consists of the whole miR-17~92a cluster with a 50 nucleotide overhang on the 3' arm. First step spliceosome factors, including SF3B1 and the spliceosome associated factor ISY1, as well as proteins involved in polyadenylation CPSF73, were shown to be involved in the generation of pro-miR-17~92²¹⁹. This work indicates that splicing does play a role in miRNA biogenesis, and the knockdown of SF3B1 leads to a reduction of mature miRNA. In addition, pri-miR-17~92a transcription is terminated by the Microprocessor²⁰⁶.

The miR-17~92a family of miRNAs plays an important role in the cell cycle, cell proliferation and apoptosis²²⁰. Knock out experiments in mice have shown that the family has an important role in the proper development of the heart and lungs, as well as a role in normal B cell development²²¹.

1.13.3: miR-23a~27a~24-2

Another example of a polycistronic miRNA genes is the miR-23a~27a~24-2 cluster which contains three miRNA hairpins (**Figure 1.21**), all of which are involved in T cell differentiation. While all three miRNAs can regulate the same processes, they can also act antagonistically to each other indicating that the regulation of the mature expression is tightly regulated²²².

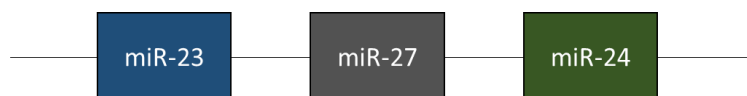


Figure 1.21: The order of the hairpins in pri-miR-23a~27a~24-2.

The miR-23a~27a~24-2 cluster is expressed from an intergenic ncRNA^{223,224}. It has a homolog, miR-23b~27b~24-1, which is located within an intron of a protein-coding gene, *AP-O* and has a role during hypoxia^{228,230}. The mature forms of miR-24-1 and miR-24-2 are identical,

while miR-23a and miR-27a differ by one nt to their homologs. The miRNA within the intergenic cluster are not necessarily expressed equally, and this may depend on the biological situation of the cell²²³.

While some lnc-pri-miRNAs are known to terminate transcription using the Microprocessor, it seems unlikely that pri-miR-23a~27a~24-2 is part of this group. When Huh7 cells were treated with DGCR8 siRNA, only a small increase in transcriptional read through was detected²⁰⁶.

1.13.4: miR-29a~29b-1

The miR-29 family consists of three mature miRNAs which are encoded by two primary transcripts (**Figure 1.22**). One transcript contains miR-29b and miR-29a and is located within an intron of a lncRNA. The second transcript also expresses miR-29b, as well as miR-29c, and this miRNA cluster is located within an exon of a nc-RNA^{226,227}. The seed regions of all three mature miR-29 are the same and are predicted to target similar genes.

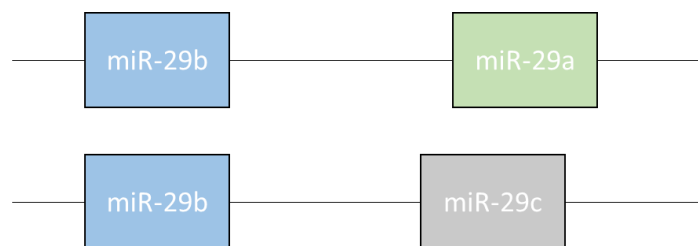


Figure 1.22: The order of miR-29 family hairpins.

While the seed regions are identical, the 3' ends do vary and has been shown to have a role in the subcellular localisation of the miR-29b family; miR-29a mainly in cytoplasm while miR-29b is enriched in the nucleus. Different cellular localisations of miRNAs may lead to a different subset of targets being recognised, and therefore, may lead to different biological roles for the miRNAs of the miR-29 family²²⁷.

Like pri-miR-122 and pri-miR-17~92a, the transcriptional termination of pri-miR-29a~29b-1 is mediated by the Microprocessor with Microprocessor knockdown lead to a 23 x increase in 3' read through²⁰⁶.

1.14: Aims and objectives

Previous research has shown that both splicing and Microprocessor cleavage occur co-transcriptionally, and while splicing and Microprocessing have been shown to interact on pre-mRNAs with intronic miRNAs, very little is known about how these two processes may interact when a miRNA is located within an exon or a lncRNA. This is especially true in respect to miRNAs located within an exon of a lncRNA, such as miR-122. The overall aim for this work is to understand the events that occur as exonic miRNAs are transcribed and the fate of the mRNA. For example, if an exonic miRNA is processed by the Microprocessor is the host mRNA still produced? Similarly, how are these two processes, miRNA biogenesis and splicing, regulated with respect to each other? This is especially important considering lncRNA biology is still poorly understood and with many miRNAs being located within lncRNAs understanding the fundamental mechanisms driving miRNA expression is important for understanding their relevance to disease.

The main focus here is miR-122 due to its clinical importance and because it is such a highly expressed miRNA it is important to understand how this high expression is achieved. By understanding the regulation behind miR-122 expression, new therapeutic drugs could be designed for miR-122 related diseases, such as HCV.

The main objectives of this work are:

- To understand how genomic location may affect biogenesis of miR-122 and the host transcript where pre-miR-122 is located within introns and exons of a β -globin gene.
- To identify what role splicing plays in the biogenesis of endogenous miR-122 in Huh7 cells and overexpressed miR-122 in HeLa cells.
- To elucidate the role of splicing and Microprocessing during the transcription of endogenous pri-miR-122.

Chapter 2: Methods

2.1: General reagents

EMD Millipore Headquarters, 290 Concord Road, Billerica, Massachusetts 01821, USA.

Bio-Rad Laboratories, 2000 Alfred Nobel Drive, Hercules, California 94547, USA.

Invitrogen, 3 Fountain Drive, Inchinnan Business Park, Paisley, PA4 9RF, UK.

Fisher-Scientific, Bishop Meadow Road, Loughborough, Leicestershire, LE11 5RG, UK.

GE Healthcare (Amersham), GE Healthcare Life Sciences, Amersham Place, Little Chalfont, Buckinghamshire, HP7 9NA, UK.

New England Biolabs (NEB), 75-77 Knowl Piece, Wilbury Way, Hitchin, Hertfordshire, SG4 0TY, UK.

Roche Products, 6 Falcon Way, Shire Park, Welwyn Garden City, AL7 1TW, UK.

Sigma-Aldrich, The Old Brickyard, New Road, Gillingham, Dorset, SP8 4XT, UK

Agilent Technologies, 5301 Stevens Creek Blvd, Santa Clara, CA95051, USA.

Table 1: Plasmids

Plasmid name	Description	Source
pLUC122x2	pGL3-MCS backbone (Promega pGL3con with MCS inserted in 3'UTR) containing two copies of the miR-122 binding region from strain H77 (1- 60nt), ie four miR-122 binding sites, inserted between SpeI and EcoRI sites in the 3'UTR of firefly luciferase.	Made by Catherine Jopling ²²⁸
Renilla	pRL-SV40	Promega
pbeta.delta.5-7	pUC18 backbone, with HIV LTR driving entire b-globin gene with 3' flanking region. Delta5-7 indicates part of 3' flanking region removed.	Gift from Nick Proudfoot ²²⁹
pHIV-LTR-pri-122	pbeta.delta.5-7 with b-globin and terminator removed and replaced with 8.5kb chunk of pri-miR-122, starting at TSS with stem-loop deleted.	Gift from Nick Proudfoot ²⁰⁶
pHIV-LTR-pri-122delta	pbeta.delta.5-7 with b-globin and terminator removed and replaced with 8.5kb chunk of pri-miR-122, starting at TSS.	Gift from Nick Proudfoot ²⁰⁶
pbeta.delta.5-7.pre122.int2	Pre-miR-122 sequence inserted within intron 2 at MfeI site of pbeta.delta.5-7	Made by Catherine Jopling
pbeta.delta.5-7.pre122deltaSD.int2	Pre-miR-122 sequence with single-double stranded region removed inserted within intron 2 at MfeI site of pbeta.delta.5-7	Made by Catherine Jopling
pbeta.delta.5-7.pre-122.ex2	Pre-miR-122 sequence inserted within exon 2 at BamHI site of pbeta.delta.5-7	Made by Catherine Jopling
pbeta.delta.5-7.pre-122delSD.ex2	Pre-miR-122 sequence with single-double stranded region removed inserted within exon 2 at BamHI site of pbeta.delta.5-7	Made by Catherine Jopling
pbeta.delta5-7.pre122.ex3	Pre-miR-122 sequence inserted within exon 3 by blunt cloning of pbeta.delta.5-7	Made by Catherine Jopling
pbeta.delta5-7.pre122dSD.ex3	Pre-miR-122 sequence with single-double stranded region removed inserted within exon 3 by blunt cloning of pbeta.delta.5-7	Made by Catherine Jopling
pbeta.delta5-7.pre122.3'FAR	Pre-miR-122 sequence inserted within the 3' flanking region by blunt cloning of pbeta.delta.5-7	Made by Catherine Jopling
pbeta.delta5-7.pre122dSD.3'FAR	Pre-miR-122 sequence with single-double stranded region removed inserted within the 3' flanking	Made by Catherine Jopling

	region by blunt cloning of pbeta.delta.5-7	
pTAT	Co-transfect with A70 to stimulate HIV promoter transcription.	Gift from Nick Proudfoot ²³⁰
pCR2.1actin	Plasmid containing actin used to make the northern blot probe.	Gift from Eric Jan

Table 2: Antibodies

Antibody	Product code	Supplier
Anti-lamin A/C	4C11	Sigma
Anti-beta-tubulin	AB6046	Abcam
Anti-Drosha	AB12286	Abcam
Anti-DGCR8	10996-1-AP	Proteintech
Anti-H2B	AB1790	Abcam
Anti-mouse	P0447	Dako
Anti-Rabbit	A6154	Sigma

Table 3: siRNA sequences

siRNA	Sequence	Position	Supplier
Drosha	cgaguaggcuucgugacuudTdT	1672-1690 in NM_013235	Dharmacon
DGCR8	caucggacaagagugugau.dT.dT	2558-2576 in NM_022720	Dharmacon
SF3b1	Proprietary	NM_001308824	Dharmacon
Control	Proprietary	N/A	Dharmacon

2.2: Tissue culture

2.2.1: Tissue culture reagents

Phosphate buffered saline (PBS): 4.3 mM Na₂HPO₄, 1.5 mM KH₂PO₄, 137 mM NaCl, 2.7 mM KCl, pH 7.4

Dulbecco's modified eagle medium (DMEM): Sigma-Aldrich

Foetal bovine serum (FBS): Sigma

Non-essential amino acids (NEAA): Invitrogen

Trypsin-EDTA (Trypsin): 0.05 % trypsin, 0.53 mM EDTA

2.2.2: Cell lines

2.2.2.1 HeLa

HeLa cells were maintained in DMEM supplemented with 10 % FBS. Cells were passaged when at 100 % confluency, typically every 2-3 days using trypsin and seeded at a density of 8.4×10^5 cells to maintain the cell line. HeLa cell lines were grown at 37 °C with 5 % CO₂.

2.2.2.2 Huh7, Huh 7.5 derived cells

Huh7 cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10 % FBS, with an addition of 1 % non-essential amino acids (Invitrogen) and 1 % L-glutamine. Cells were passaged as outlined in 3.1.1 using a seeding density of 1.4×10^6 cells. Cells were grown at 37 °C with 5 % CO₂.

2.2.3 Cell maintenance

For routine passage, cells were grown in a T75 flask until confluency, then washed twice in PBS. 2 ml of trypsin was added to the flask, which was incubated for 2-10 minutes at 37 °C with 5 % CO₂ until the cells have dissociated from the flask. To neutralise the trypsin, fresh pre-warmed DMEM was added to the flask and the cell density altered to seeding confluency. For specific seeding densities, a haemocytometer was used to count the cells. Cell lines are discarded after approximately 20 passages.

2.2.4 Cryopreservation

To preserve the cell lines, cells were cryopreserved in liquid nitrogen following thawing. After the first passage, 2.2×10^6 cells are seeded into a 10 cm plate. When at confluency, approximately 8.8×10^6 cells, the cells are trypsinised and resuspended in 1 ml of 90 % FBS with 10 % DMSO. For maximum recovery of cells, vials for cryopreservation were stored at -80 °C before being transferred to liquid nitrogen for long term storage.

2.2.5: Cell Thawing

Following removal of liquid nitrogen storage, cryopreserved cells are rapidly thawed in a 37 °C water bath. The cells are then diluted in pre-warmed media appropriate to the cell line. The cells are then pelleted and the supernatant discarded to remove the DMSO used in cryopreservation.

The cell pellet is gently resuspended in the appropriate medium for the cell line and cultured as outlined in 2.1.2.1 or 2.1.2.2.

2.3: Bacterial methods

2.3.1: Reagents and solutions

Lysogeny Broth (LB) medium: 1 % tryptone; 0.5 % yeast extract; 0.5 % NaCl; pH adjusted to 7.5 with NaOH then autoclaved.

Lysogeny broth (LB) agar: 1.5 % agar, 1 % tryptone; 0.5 % yeast extract; 0.5 % NaCl; pH adjusted to 7.5 with NaOH then autoclaved. Once cooled, 100 mg/ml ampicillin added prior to pouring into plates.

2.3.2: Bacterial strains

DH5α cells were obtained from NEB and have the following genotype: fhuA2 Δ (argF-lacZ) U169 phoA glnV44 Φ80 Δ (lacZ) M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17.

2.3.3: Preparation of competent cells

DH5α cells were prepared by Barbara Rampersad (Gene Regulation and RNA Biology laboratory technician, University of Nottingham).

2.3.4: Transformation

1 ng of plasmid was mixed with 50µl of DH5α and incubated on ice for 5-10 minutes. Cells were then heat-shocked at 42 °C for 60 seconds. Heat-shocked cells were then incubated at 37 °C for 1 hour with shaking with 150µl of LB. The transformed cells were spread onto agar plates and incubated overnight at 37 °C. An individual colony can then be used to inoculate 5 ml of LB with antibiotic, which was incubated overnight at 37 °C.

2.3.5: Miniprep

Following overnight incubation as outlined in 2.2.4, the cells were centrifuged at 4000 rpm to pellet the cells. Plasmid DNA was extracted by mini-prep using the Wizard Plus SV Minipreps DNA purification kit (Promega) following the manufacturers protocol.

2.3.6: Maxiprep

When a greater quantity of plasmid is required, a maxiprep was used instead of a miniprep. 100µl of culture, described in 2.2.4, was inoculated into 100ml of LB with antibiotic and grown at 37 °C overnight with shaking. The next morning, cells were pelleted by centrifugation at 4000 rpm for 15 minutes at 4 °C. Plasmid DNA was extracted by maxi-prep using a NucleoSpin maxiprep kit (Machery-Nagel) following the manufacturer's protocol.

2.4: Plasmid Transfection

Plasmid transfections were performed using Lipofectamine 2000 (Invitrogen). 6 well plates were seeded with 1.78×10^5 of HeLa cells per well 24 hours prior to transfection to ensure the cells were at 70 % confluency. Per well, 500 µl of OptiMem was incubated for 5 minutes at room temperature with 1.25 µl Lipofectamine 2000. This was then incubated for 20 minutes at room temperature with 500 µl of OptiMem containing 0.1 µg of plasmid and 0.025 µg of pTAT. The cell media was removed and replaced with 1 ml of transfection mix per well. After 6 hours, the transfection mix was removed and replaced with fresh pre-warmed media. The cells were harvested 24 hours following the addition of transfection mix using Tri reagent, unless further treatment was necessary.

2.5: siRNA knockdown

To knockdown a protein of interest, cells were transfected with siRNA using Lipofectamine RNAiMax (Invitrogen). Cells were seeded in a 6 well plate 24 hours before transfecting so that the confluency starting transfection was 30 %. For each well, 150 µl of OptiMem was incubated for 5 minutes at room temperature with 1.5 µl Lipofectamine RNAiMax and then incubated for a further 20 minutes with 150 µl of Optimem containing 20 nM of siRNA. The cell media was then removed and replaced with 300 µl of transfection mix and 700 µl of pre-warmed media. After 6 hours, the media was replaced with fresh pre-warmed media. A second siRNA knockdown was repeated 48 hours after the first knockdown. The cells were harvested in Tri reagent unless being treated further.

2.6: Morpholino transfection

Morpholinos (AMOs) were transfected into Huh7 using the Neon transfection system (Invitrogen). Cells were grown to 70 % confluency,

trypsinised and resuspended in fresh media before being counted using a haemocytometer. 4×10^5 cells were pelleted per transfection at 1100 rpm for 5 min. The pelleted cells were washed with Ca^{2+} and Mg^{2+} free PBS (Invitrogen) and resuspended into 10 μl of buffer R and 100 nM of AMO added. Electroporation was performed using programme 15: 2 pulses, 1300 V, 20 width. The cells were added to fresh media and plated into a 6 well plate. Cells were harvested using Tri reagent.

2.7: RNA isolation

Prior to RNA extraction, cells were washed twice with chilled PBS. Cells were then incubated for 5 minutes in Tri reagent (Sigma), with 1 ml of Tri reagent used for 10^7 cells. The lysed cells were then transferred to a microfuge tube (volumes larger than 1 ml were aliquoted and extracted in parallel). 200 μl chloroform was added to each tube and vigorously shaken for 30 seconds, incubated at RT for 2 minutes before being centrifuged for 15 minutes, at 4 °C and 13000 rpm. The aqueous phase was incubated at -20 °C for at least 10 minutes (1 μl glycoblue (Invitrogen) was added if yields are expected to be low). The precipitated RNA was then pelleted by centrifugation (4 °C, 10 minutes, 13000 rpm). The supernatant is then discarded and the pellet washed in 70 % ethanol. Following centrifugation (4 °C, 5 minutes, 13000 rpm) the pellet is allowed to dry before being resuspended in 40 μl RNase-free water.

2.8: DNase treatment

For pri-miRNA isolation, the RNA pellet from section 2.4 was resuspended in 200 μl DNase I buffer (10 mM Tris-HCl pH 7.5, 10 mM MgCl_2). 2 μl of RQ1 RNase-free DNase (Promega) was added to the RNA samples and incubated at 37 °C for 1 hour before undergoing an acid-phenol chloroform extraction. An equal volume of acid-phenol chloroform was added to the DNase treated sample. The mixture was vigorously shaken for 30 seconds, rested for 2 minutes before being centrifuged for 5 minutes at 4 °C and 13000 rpm. The RNA containing aqueous phase was transferred into a new microfuge tube, and an equal volume of isopropanol added, plus 10 % 5 M NaCl (and glycoblue (Invitrogen) if required). The RNA was precipitated for 10 minutes at -20 °C before being centrifuged for 10 minutes at 4 °C and 13000 rpm. The RNA pellet was then wash in 70 % ethanol and centrifuged

for 5 minutes at 4 °C and 13000 rpm. The pellet was allowed to dry before being resuspended in 20-40 µl RNase-free water.

2.9: Determination of nucleic acid concentration

The concentration of DNA and RNA were determined by using a Nanodrop spectrophotometer (Nanodrop ND-1000, Thermo Fisher Scientific) by measuring the absorbance at 260 nm.

2.10: mRNA- Reverse transcription

To analyse the levels of specific mRNAs by RT-qPCR, purified RNA first has to be reverse transcribed. The first step is to add 150 ng of random primers (Invitrogen) and 10 mM dNTPs (Invitrogen) to 100 ng of total RNA to a total volume of 14.25 µl. This mix was then heated to 65 °C for 5 minutes before cooling on ice. To this, 5.75 µl of a master mix containing 5 X SuperScript III buffer, 0.1 M DTT, 100 U SSIII and 5 U RNasin (all Invitrogen) was added. The reaction was incubated for 2 min at RT, then 1 hr at 50 °C before heating at 70 °C for 15 min to deactivate the enzyme.

2.11: mRNA- qPCR

Real-time quantitative PCR (qPCR) was performed using a Rotor-gene Q 2lex qPCR machine (Qiagen) using the following cycle parameters: 95 °C for 10 min, 40 cycles of 95 °C for 10 sec, 58 °C for 15 sec, 72 °C for 20 sec, and a final cycle of 95 °C for 1 minute, 58 °C for 30 sec, 95 °C for 30 sec. For the qPCR reaction, 0.5 µl of cDNA produced as outlined in 2.6 was added to 2 X GoTaq SYBR Green master mix (Promega), and 10 µM of forward and reverse primer to a final volume of 10 µl. Each reaction was performed in triplicate.

2.12: miRNA- Reverse transcription

For mature miRNA, reverse transcription was carried out using a TaqMan MiRNA Reverse Transcription kit using a miRNA-specific RT primer (all Applied Biosciences). 100 ng of total RNA, obtained as described in 2.5, was added to a master mix containing: 1.5 µM dNTPs, 1 U RNase Inhibitor, 10 X RT buffer, 2.5 U MultiScribe Reverse Transcriptase, and 20 X specific miRNA RT primer to a final volume of 7.25 µl. The reaction was incubated in a thermocycler with the cycle: 16 °C for 30 minutes, 42 °C for 30 minutes, 85 °C for 5 minutes.

2.13: miRNA- qPCR

qPCR was performed in triplicate in a 10 µl reaction using a final concentration of 1 x Taqman No Amperase UNG mix (Invitrogen), 0.5 µl of the miRNA-specific probe and 0.665 µl of cDNA, using a Statagene Mx3005p qPCR machine. The following qPCR cycle was used: 95 °C for 10 minutes, and 40 cycles of 95 °C for 15 sec, 60 °C for 1 minute.

Table 4: A table of qPCR primers used.

Type of primer	Name	Sequence
Control	Actin qF	agcacagagcctcgcttt
	Actin qR	tcatcatccatggtgagctg
Pri-miR-122	pri122 int qF (Unspliced)	gggatcctcctgcctaagac
	pri122 int qR (Unspliced)	gaaaaggccttgagggattc
	pri122 spl qF	ctgactggggattattcaaagc
	pri122 spl qR	tcgcgaccatccacaagag
	Hairpin qF	cccgtgatgcttctttctc
	Hairpin qR	ccattgtcacactccacagc
	3' miR122 qF	cccaggcttggttaaataag
	3'miR122 qR	acagtgtcaggcacacggta
Pri-miR-17~92	Pri-17 qF	ggaagccagaagaggagga
	Pri-17 qR	aagtgggtggctctccaatc
Pri-miR-23a~24-2~27a	Pri-23 qF	gcctggagagaagggatctt
	Pri-23 qR	agggactccagagatgacca
Inc-pri-miR-29a-29b-1	Pri-29 qF	atggagcagcacttgcttgctt
	Pri-29 qR	gtcactcttggtgggaaga
β-globin	β-globin int qF	ccaaatcagggttaatttgcatt
	β-globin int qR	tgctattgccttaaccagaa
	β-globin ex qF	gaaggctcatggcaagaaag
	β-globin ex qR	cactggtgggtgaattctt

2.14: Analysis of qPCR data

All qPCR data was analysed using the $2^{-\Delta\Delta C_t}$ method¹⁰³. To determine the statistical significance between two sets of samples, a t-test was used and a one-way ANOVA for larger data sets. For comparisons between two sets of related data, for example, treated and untreated, a two-way ANOVA was used. A ROUT analysis was used to identify outliers.

2.15: Northern blot for mRNA

2.15.1: Solutions

Northern mix: 1 ml formamide, 350 µl formaldehyde, 200 µl

Gel loading buffer: 50% glycerol, 1 mM EDTA, 0.4 % bromophenol blue, 0.4 % xylene cyanol

10 x MOPS buffer: 0.2 M MOPS, 0.05 M NaAc, 10 mM EDTA pH 8.0, pH to 7.0 with NaOH. Store in the dark.

20 x SSC: 3 M NaCl, 0.3 M sodium citrate, pH 7.0.

Church-Gilbert's hybridization buffer: 1mM EDTA, 1% BSA, 0.5M sodium phosphate monobasic, 7% SDS, pH 7.2.

Wash 1: 2 x SSC, 0.1 % SDS

Wash 2: 0.5 x SSC, 0.1 % SDS

Wash 3: 0.01 x SSC, 0.1 % SDS

Stripping buffer: 0.5 % SDS in boiling water

2.15.2 RNA agarose gel electrophoresis

Northern blotting for mRNA was performed using 1 µg of total RNA denatured for 10 minutes at 65 °C with 15.5 µl northern mix. 2 µl of gel loading buffer was added to the samples before being loaded on to the gel. The gel was made by heating 1 g of agarose in 73 ml of RNase-free water until the agarose had dissolved. The heated agarose was allowed to cool before 1 x MOPS buffer and 17 % formaldehyde were added to a total volume of 100 ml. with a running buffer containing 1 X MOPS and 10 % formaldehyde. The the gel was run at 100V for 1-2 hours until the bromophenol blue dye had migrated ~2/3 of the length of the gel.

2.15.3: Blotting

Once the RNA was sufficiently separated, the gel was soaked for 20 minutes in 0.05 M NaOH before soaking for 40 minutes in 20 X SSC. To create the transfer tower (**Figure 2.1**), a dish was filled with 20 X SSC and a glass plate was laid on top. A wick made of filter paper slightly wider than the size of the gel was laid on the glass plates and the ends submerged in the 20 X SSC solution. The gel was placed on the wick, followed by three pieces of blotting paper, and the Zeta Probe membrane (Bio-Rad). After the addition of each layer, trapped air was removed by gently rolling with a pipette. A stack of tissue was placed on top of the membrane, followed by a second glass plate and weights to compress the tower. Parafilm was placed around the sides of the wick to prevent to block

any direct transfer of 20 X SSC from the wick to the tissue. The tower was left overnight.

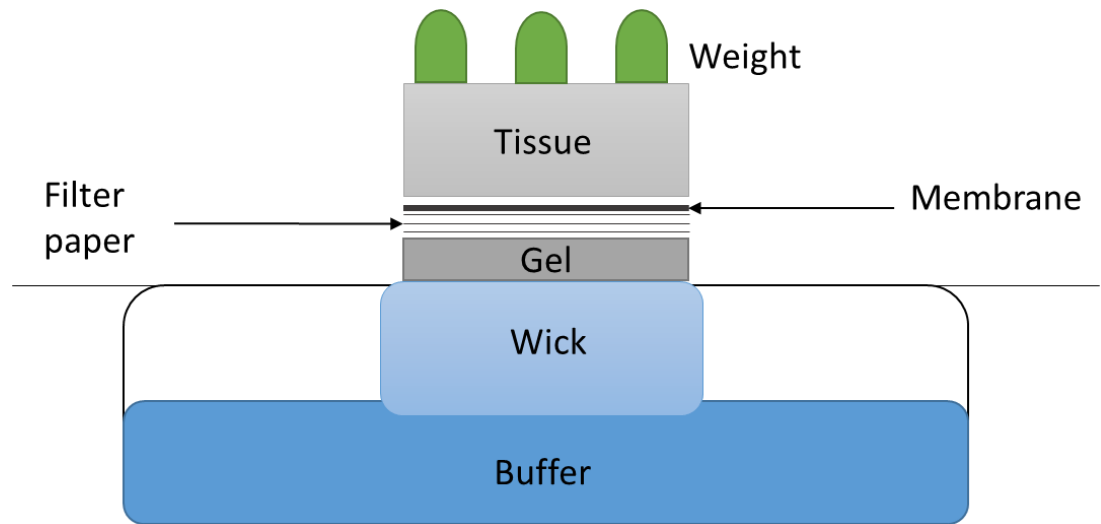


Figure 2.1: A diagram of the northern blotting tower

2.15.4: Crosslinking and hybridisation

The following day, the tower was disassembled. The membrane was crosslinked using a Stratalinker 2400 (Stratagene) at 1200 mJ/min. Positions of ribosomal RNA were identified and RNA integrity confirmed using methylene blue staining- the membrane was incubated in methylene blue solution for 2 min and then rinsed with RNase-free water until the non-bound dye was removed. The two bands corresponding to the ribosomal RNA was marked. Membranes were pre-hybridised in Church Gilberts solution at 65 °C for 30 minutes, before incubating with 12.5-15 mg radiolabelled probe overnight. The blot is then washed for 10 minutes at room temperature first with wash buffer 1, then wash buffer 2 and 3. Once washed, the blot is rinsed with deionised water, gently dried and wrapped in Saran wrap. The blot is then placed on a phosphorimager sheet and incubated overnight before imaging on a Storm 825 phosphorimager.

2.15.5: Probe preparation

The probe was prepared by boiling 30 mg of DNA in a volume of 15 µl for 3 minutes before snap cooling. DNA was incubated for 1-2 hours at RT with 5 x labelling buffer (Promega), 250 nM dNTP mix (A, G and T only), 400 ng BSA, 5 U Klenow exo- and 2.5 µl α-dCTP. The membrane was imaged using a Storm phosphorimager.

The fragments used to create the probes were obtained by restriction enzyme digestion. For actin, a plasmid containing actin was digested with EcoRI for 1 hour at 37 °C. To create the β -globin fragment, the β -globin empty vector plasmid was digested with BglII and BamHI for 1 hour at 37 °C. Once the fragments were digested, the DNA was separated on a 1 % agarose gel as outlined in **2.24** and the appropriate size band was cut out and gel extracted as outlined in **2.25**.

2.15.6: Stripping

To remove bound radioactive label, the blot is washed for 30 minutes three times in 0.5 % SDS in boiling water. The blot can then be rinsed with deionised water and imaged as described above.

2.16 Northern blotting for mature miRNA

2.16.1: Solutions

50 x MOPS: 1 M MOPS, pH7.

dPAG: 7 M urea, 15 % bis-acrylamide (40 %), 1 x MOPS. 50 ml is required per gel, and set with 100 μ l 10 % APS and 50 μ l TEMED.

Loading dye: 95 % formamide, 1 mM EDTA-RNase free, 0.4 % bromophenol blue, 0.4 % xylene cyanol

EDC solution: 0.16 M EDC, 0.13 M 1-methylimidazole, pH 8 with HCl.

Pre-hybridisation buffer: 5 x SSPE, 6.67 x Denhardts, 0.1 X SDS

50 x Denhardts: 1 % ficol, 1 % polyvinylpyrrolidone, 1 % bovine serum albumin.

Washing buffer: 2 x SSC, 0.5 % SDS

20 x SSC: 3 M NaCl, 0.3 sodium citrate, pH 7.0.

Stripping buffer: 0.1 x SSPE, 0.1 % SDS

20 x SSPE: 3 M NaCl, 0.2 M NaH_2PO_4 , 0.02 M EDTA, pH 7.4 and autoclaved.

2.16.2: miRNA agarose gel electrophoresis

For mature miRNA, northern blotting protocol was followed as described in the literature until probing²³¹. To make the gel, the urea was dissolved in RNase-free water, acrylamide and MOPs buffer as indicated as above in **2.12.1**. 3 ml of the dPAG solution was used to create a plug, with 100 μ l 10 % APS and 50 μ l TEMED added to ensure the plug set quickly to prevent leaks. The same quantities of TEMED and APS were added to the

remaining solution and poured on top of the plug. The wells were washed with running buffer to remove settled acrylamide and pieces of gel before pre-warming at 18 W for 30 minutes. In this time, 10 µg of total RNA was denatured by heating at 75 °C for 10 minutes with loading dye and chilled. Once the gel was warmed, the wells were again washed with running buffer, and the RNA samples loaded. The RNA was separated on the dPAG gel at 18 W for 90 min using 1 x MOPS running buffer.

2.16.3 Transfer and hybridisation

RNA was transferred onto HyBond NX membrane (Amersham) using a semi-dry transfer machine for set at 10 v for 1 hour at 4 °C. The blot was then heated at 60 °C with EDC solution to crosslink the RNA for 1 hour. The blot was incubated at 37 °C for 30 minutes with 7 ml of pre-warmed ULTRAhyb-oligo hybridisation buffer (ThermoFisher) or the hybridisation buffer described in 2.12.1 before the addition of 25 µl of radiolabelled probe which was left to incubate overnight. Once probed, the blot was washed twice in washing buffer for 30 minutes at 37 °C. The blot was gently dried before being wrapped in Saran wrap. The blot was placed on a phosphorimager screen overnight before imaging with a Storm 825 phosphorimager.

2.16.3: Stripping

To remove bound radiolabelled oligo, the blot was washed for 30 minutes three times in stripping buffer at 65 °C. The blot was then rinsed with deionised water and gently dried before being placed on a phosphorimager screen as described in 2.11.1.

2.16.4: Probe preparation

In a 10 µl reaction, 1 µl of anti-sense oligo (100 µM), 1 µl PNK buffer, 2.5 µl γ-ATP and 1 µl T4 PNK were incubated for 30 minutes at 37 °C. 40 µl of RNase-free water was used to dilute the probe. Unincorporated nucleotides were removed using an Illustra MicroSpin G-50 column (Life Sciences). Anti-sense oligos for U6 (ATATGGAACGCTTCACGAATT) and miR-122 (ACAAACACCATTGTCACACTCCA) were used.

2.17: Splice inhibition

To inhibit splicing, cells were treated with Pladienolide B (Pla B; Sigma) dissolved in DMSO. Cells were treated for 4 hours before harvesting (**2.3 or 2.7**) with 1 μ M Pla B, or DMSO, unless otherwise specified.

2.18: AlamarBlue cell viability assay

AlamarBlue (ThermoFisher) was used following the manufacture's guidelines. Cells were treated at 70 % confluency in 96 well plates and were treated in replicates of 3-6. Cells were incubated for 1, 4, or 24 hours with the drug of interest diluted in the appropriate media, and 10 % AlamarBlue added in addition. The absorbance was measured using a BioTek Synergy HT plate reader.

2.19: Chromatin-associated RNA extraction

Cells were grown in 10 cm plates until 80 % confluent and washed twice in ice-cold PBS. The cells were then incubated on ice for 5 minutes with 4 ml ice-cold HLP1 (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 0.5 % NP-40, 2.5 mM $MgCl_2$) and the lysate was underlayered with 1 ml HLP2 (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 0.5 % NP-40, 2.5 mM $MgCl_2$ and 10 % glycerol) and centrifuged for 5 mins (4 °C at 1000 rpm). The cytoplasmic fraction was removed and the nuclei pellet suspended in 125 μ l NUN1 (20 mM Tris-HCl pH 8.0, 75 mM NaCl, 0.5 mM RNase-free EDTA, 50 % glycerol) supplemented with protease inhibitor. The resuspended nuclei were incubated on ice for 15 minutes with 1.2 ml of NUN2 (20 mM Hepes, 7.5 mM NaCl, 0.2 mM RNase-free EDTA, 1 M urea, 1 % NP-40) supplemented with protease inhibitor and vortexed regularly before centrifugation for 10 minutes at 13000 rpm at 4 °C. The supernatant containing the nucleoplasmic fraction was removed and the chromatin pellet suspended in 200 μ l HSB (10 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM $MgCl_2$) and treated with Turbo DNase and Proteinase K for 10 minutes at 37 °C. All fractions were either extracted in Tri Reagent or acid-phenol chloroform. Protein samples were taken from the supernatant, or from the resuspended pellet with 1/250 taken from the cytoplasmic fraction and 1/25 taken from the nucleoplasm and chromatin fractions.

2.20: Luciferase assay

HeLa cells were seeded in a 24 well plate to a starting density of 3.4×10^4 per well prior to transfecting (**2.2**) with the plasmid of interest including

pTAT. Cells were co-transfected with a pGL3-mcs firefly luciferase reporter vector and the control renilla luciferase plasmid. Cells were harvested using 5 x Passive Lysis Buffer (Promega), incubated for 15 minutes at room temperature while rocking. Renilla and firefly luciferase activity was analysed using the Dual-luciferase reporter assay system (Promega) following the manufacturers guidelines. Each assay was measured in triplicate, and the average of these used in the analysis.

2.21: SDS-PAGE

Protein was extracted from cells, unless otherwise specified, by incubating in RIPA buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1.0% NP-40, 0.5% sodium deoxycholate, and 0.1% sodium dodecyl sulphate) for 5 minutes on ice. Cells were then scraped, and cell debris pelleted. Protein concentration was analysed using a Bradford assay. Samples were boiled for 10 minutes before loading onto a 12 % SDS-PAGE gel and run for 90 minutes at 180 V in running buffer (25 mM Tris-HCl pH 8.3, 192 mM glycine, 0.1 % SDS).

Table 5: SDS-PAGE

Resolving gel	8 %	12%	Stacking gel	
30 % acrylamide/ 0.8 % bis-acrylamide	4.0 ml	6 ml	30 % acrylamide/ 0.8 % bis- acrylamide	650 µl
1.5 M Tris-HCl pH 8.8	3.8 ml	3.75 ml	1 M Tris-HCl pH 6.8	625 µl
10 % SDS	150 µl	150 µl	10 % SDS	50 µl
ddH2O	6.9 ml	5.1 ml	ddH2O	3.625 ml
10 % APS	100 µl	100 µl	10 % APS	50 µl
TEMED	20 µl	20 µl	TEMED	10 µl

2.22: Bradford assay

To ensure the equal loading of protein samples for SDS-PAGE electrophoresis, a Bradford assay was used to determine the concentrations of proteins. 2 µl of protein sample was added to 1 X Bradford reagent (Bio-Rad). A standard curve was created using 1 X Bradford reagent, 2 µl of the appropriate buffer, and 2-10 µg of BSA. The samples were incubated for 30 min at RT and the absorbance measured at

595 nm using an Eppendorf photometer. The absorbance of the sample was compared to the BSA standards to determine protein concentration.

2.23: Western blotting

Protein from SDS-PAGE gels were transferred on to a PVDF membrane (Amersham) pre-soaked in methanol using a semi-dry transfer machine. Blotting paper was soaked in transfer buffer (50 mM Tris-HCl pH 8.3, 192 mM glycine, 20 % methanol). The proteins were transferred for 1-1.5 hours at 70 mA at room temperature. The membrane was then blocked for 1 hour in 5 % milk solution in TBST (10 mM Tris-HCl pH 8.0, 0.9 % NaCl, 0.1 % Tween-20) before overnight incubation at 4 °C in primary antibody. After washing in TBST, the membrane was then incubated at room temperature in the appropriate secondary antibody for 1 hour. The blot was imaged using Amersham ECL Prime Western Blotting Detection Reagent (GE Life Sciences) and ImageQuant LAS-500.

2.24: DNA agarose gel electrophoresis

1 % agarose gels were used to separate nucleic acids according to molecular weight by electrophoresis. 1 g of agarose was dissolved in 1 X TAE. Once cooled SYBR-safe (Invitrogen) was added and poured into a gel cast. Samples containing 6 X loading dye (30% (v/v) glycerol, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol were loaded on to the gel and the gel was run at 100 V for 30 min to 1 hour. Nucleic acids were visualized with a UV transilluminator or an ImageQuant LAS-500.

2.25: Gel extraction

Following a restriction enzyme digestion, bands of interest were excised from the gel for gel extraction using a Monarch gel extraction kit (NEB).

Chapter 3: Intronic and exonic miR-122 expression in HeLa

3.1: Introduction

With both the spliceosome and Microprocessor known to interact with the carboxyl-terminal domain of RNA polymerase II^{11,12,13,71} it could be theorised that these two co-transcriptional processes interact. It is known that when a miRNA is located within an intron, both miRNA and spliced transcript can be produced¹⁰², but little is known about how splicing and miRNA biogenesis may interact when the miRNA is located within an exon. Interestingly, there is an example of a switch which regulates the exonic miR-198 expression and the protein expression of its host gene *FSTL1*. KHSRP, a splicing regulatory protein, recognises a GUG domain in the pre-miRNA and induces Microprocessing, while miR-198 failed to be expressed when KHSRP levels were reduced, which lead to an increase of the *FSTL1* transcript being exported to the cytoplasm¹⁹⁷.

As yet, there has been no comparisons looking at miRNA biogenesis from intronic or exonic locations, or from a lncRNA context. In the most recent report looking at the genomic location of miRNA hairpins in humans, 75 % of miRNA hairpins have been shown to be located within protein coding genes are intronic¹⁰³, where previously it was estimated to be 81 %¹⁰². In comparison, just 39 % of miRNA are intronic when looking purely at hairpins located within lncRNAs¹⁰³.

To understand whether there are differences in miRNA expression of miRNAs from an intronic or exonic location, a well-defined β -globin reporter construct containing a HIV-LTR promoter with three exons and two introns was used²²⁹. The β -globin plasmids were co-transfected with a second plasmid encoding the HIV TAT transcriptional activator to enhance expression²³³. To test the influence of intron or exonic location on miRNA biogenesis, pre-miR-122 hairpin and the 65 nt downstream and 39 nt upstream relative from nt 1 of the mature strand was inserted within difference regions of the β -globin transcript (**Figure 3.1**). HeLa cells, which do not endogenously express miR-122 or β -globin, were used to express the pre-miR-122 containing β -globin plasmids. By analysing the levels of mature miR-122 produced, or the levels of spliced and unspliced host β -globin transcript, the effect of genomic location on miRNA

biogenesis and of miRNA processing on splicing can be elucidated.

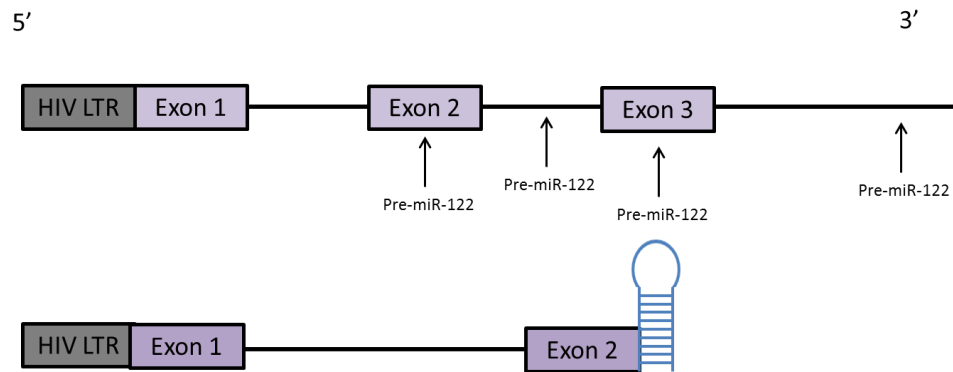


Figure 3.1: Schematic of miR-122 reporter constructs. A) β -globin reporter under the control of a HIV LTR promoter showing the possible locations of wild-type and Drosha-cleavage impaired pre-miR-122. **B)** Full length pri-miR-122 IncRNA under the same HIV LTR reporter.

Pre-miR-122 inserted within exons 2 and 3, and intron 2 as well at a site labelled 3', which is past the polyadenylation site (PAS) and therefore is transcribed as a result of transcriptional read through, were transfected into HeLa cells. Into the same locations, a Drosha-cleavage impaired miR-122 hairpin was inserted which lacks the sequence upstream of the mature miR-122 sequence, as a control to determine whether any differences in expression was due to Microprocessor cleavage. Drosha cleavage impaired miR-122 will be referred to a Δ pre-miR-122 (**Figure 3.2**).

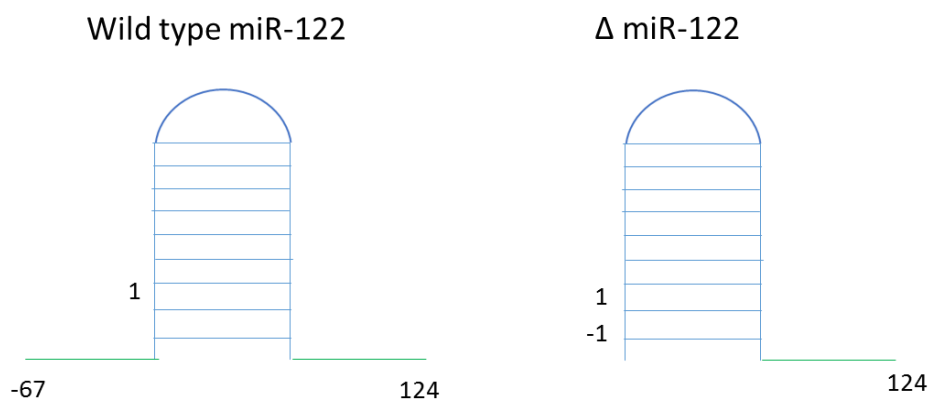


Figure 3.2: The miR-122 hairpin and the Drosha-cleavage mutant. Position 1 is the first nucleotide of the mature strand.

As a second control, the full length Inc-pri-miR-122 was inserted into the same plasmid backbone, lacking the β -globin gene but under the same HIV LTR promoter. As miR-122 is liver specific, these plasmids were transfected into HeLa cells which do not express miR-122. Therefore, in this chapter, the expression of miR-122 from its endogenous locus can be

compared to the expression of miR-122 from exonic and intronic locations. All the β -globin plasmids used in this chapter were made by Catherine Jopling and the WT and Δ pre-miR-122 plasmids were made by Ashish Dhir²⁰⁶.

3.2: miR-122 expression from intronic and exonic locations of β -globin

To compare how the genomic location affects mature miR-122 expression, total RNA was extracted from HeLa cells transfected with plasmids containing pre-miR-122 in different locations of β -globin. The levels of mature miR-122 expression was analysed using qPCR relative to a U6 snRNA control (**Figure 3.3**). The results show that mature miR-122 levels are increased when pre-miR-122 is expressed from an intronic or exonic location compared to its endogenous context as an lncRNA from a plasmid, and less mature miR-122 is expressed from the 3' flanking region compared to the lncRNA. The data also shows that there is no significant difference between the exonic and intronic expression levels, with similar levels of mature miR-122 being expressed from intron 2, exon 2 and exon 3. Cells transfected with plasmids containing Drosha-cleavage impaired pre-miR-122 fail to express mature miR-122 and it was not detected by qPCR and therefore cannot be shown in the data, confirming that miR-122 was generated by Drosha cleavage at the expected location.

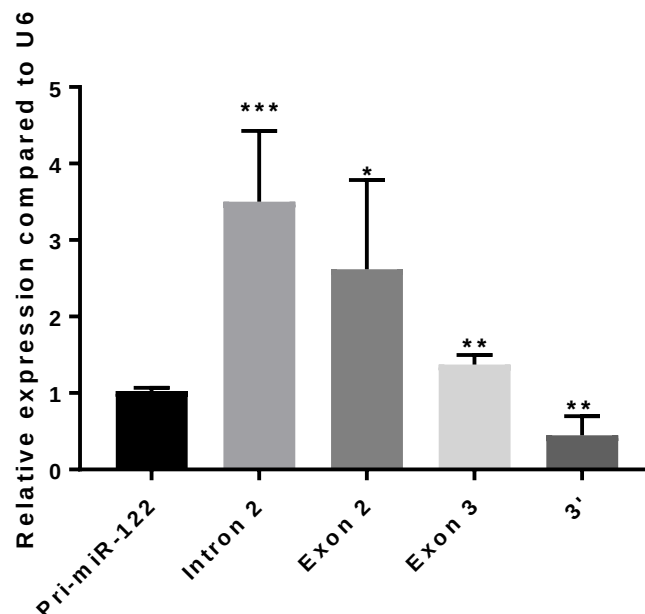


Figure 3.3: Mature miR-122 expression is increased from a genic location compared to a lncRNA. RT-qPCR analysis showing the expression of miR-122 when located with in an intron or an exon compared to endogenous expression under the same reporter. Error bars represent standard deviation with a minimum

of 2 biological replicates and p-values calculated using a t-test compared to pri-miR-122 with *p< 0.05, **p< 0.01 and ***p< 0.005.

The expression of mature miR-122 from different locations of the β -globin reporter was also analysed by northern blot. This analysis agreed with the qPCR data, showing that the expression of mature miR-122 from intronic and exonic locations is increased compared to the lnc-pri-miR-122 plasmid, and when compared to expression from the 3' flanking region (**Figure 3.4**). Again, this data shows that intronic or exonic location has little effect on miRNA expression when pre-miR-122 is expressed from these locations of a β -globin reporter, but that mature miR-122 production is enhanced when pre-miR-122 is expressed from the β -globin coding region compared to the lncRNA context.

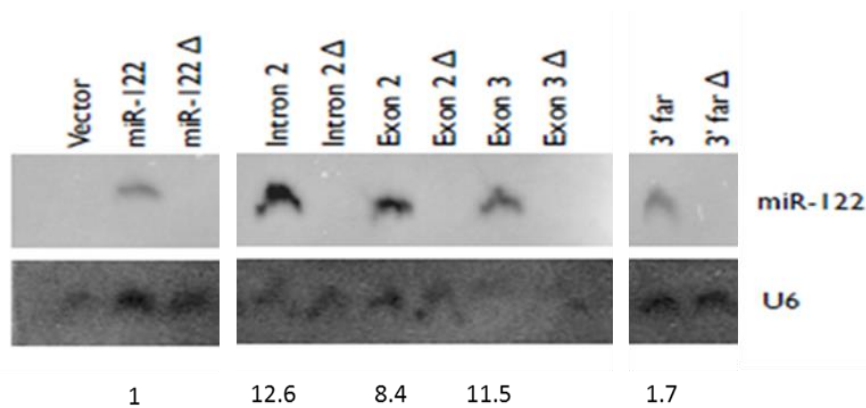


Figure 3.4: Northern blot analysis shows that mature miR-122 expression increases when located within a genic region compared to a lncRNA.

Northern blot analysis showing the expression of miR-122 when located within an intron or an exon compared to endogenous expression under the same reporter. The numbers below are quantifications of the level of mature miR-122 relative to U6 compared to miR-122.

3.3: miR-122 function when expressed from intronic or exonic locations

A luciferase reporter assay was used to determine whether there are any functional differences in miR-122 expressed from different locations of the β -globin reporter used previously. HeLa cells were transfected with the β -globin reporter used in **3.1** in conjunction with a plasmid containing a firefly luciferase construct with four miR-122 binding sites in the 3' flanking region, such that miR-122 represses luciferase expression (**Figure 3.5**). Levels of firefly luciferase activity were analysed by comparison to a renilla luciferase transfection control.

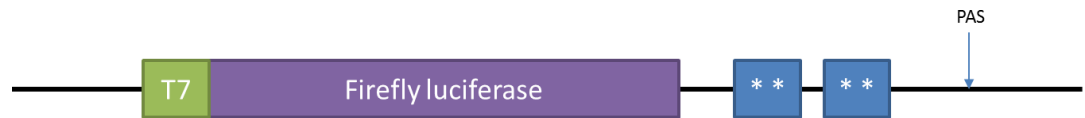


Figure 3.5: Schematic of firefly luciferase reporter. Each * represents a single miR-122 binding site.

The results show that all the β -globin plasmids which contain miR-122 reduced luciferase expression compared to the Δ pri-miR-122 control, confirming that the miR-122 expressed is functional (**Figure 3.6**). When miR-122 is expressed from the β -globin introns and exons, luciferase activity is reduced to a similar level; however, there is slightly more repression by miR-122 when expressed within the construct than from its normal lncRNA location under the same promoter. However, it is possible that there may be a small correlation between miR-122 expression levels and luciferase repression, but the lack of difference could be due to the repression machinery becoming saturated (**Figure 3.3**). This could explain why the difference between the different plasmids is less apparent given the difference in mature miR-122 expression.

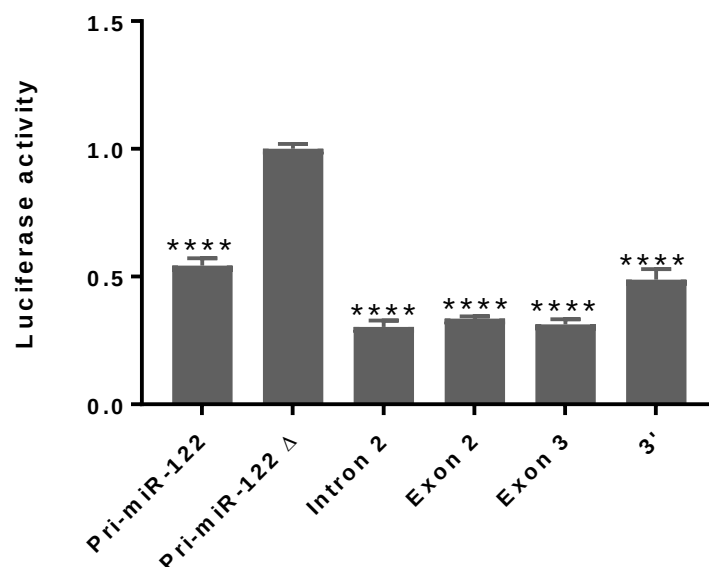


Figure 3.6: Repression of a luciferase reporter correlates with expression levels of miR-122 from different locations of a β -globin reporter. A graph showing luciferase activity relative to a renilla control. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA compared to pri-miR-122 with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

In order to quantify the levels of luciferase repression when miR-122 was expressed from plasmids in different locations, the relative expression of

luciferase activity was compared to the level of luciferase repression when known quantities of pre-miR-122 oligonucleotides were transfected; this allows for the estimation of the amount of miR-122 expressed from each site. When 20 μM to 5 μM of pre-miR-122 oligo was transfected into the cells, a similar level of luciferase activity was recorded (**Figure 3.7**). This indicates that the levels of repression reaches its maximum at 5 μM of pre-miR-122 oligo, and increased levels of oligo will have no further effect on repression. This can also be seen in figure **3.3**, where miR-122 expressed from an intronic or exonic location in the β -globin plasmid shows only slightly more repression compared to the expression of pre-miR-122 from the lncRNA control or the 3' flanking region of the β -globin reporter.

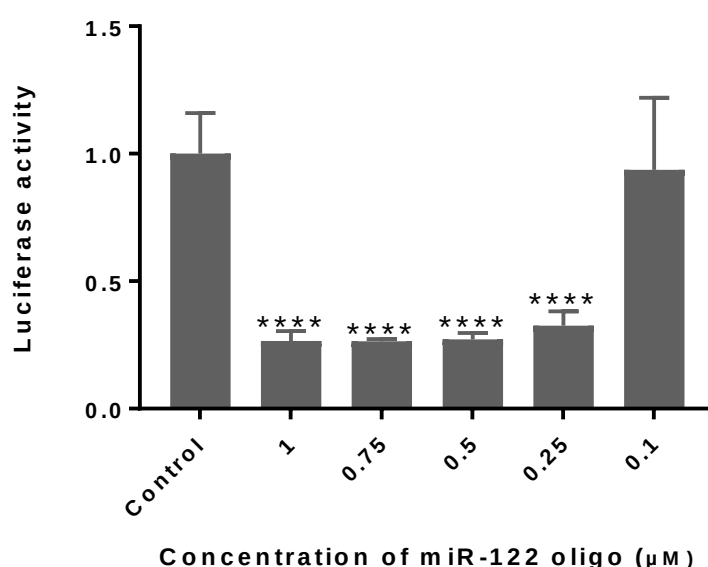


Figure 3.7: Repression of a luciferase reporter occurs with more than 0.1 μM of pre-miR-122 oligo. Graphs showing luciferase activity relative to a renilla control. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA compared to pri-miR-122 with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

Differences in luciferase activity can be seen when cells were transfected with lower concentrations of pre-miR-122 oligo (**Figure 3.8**).

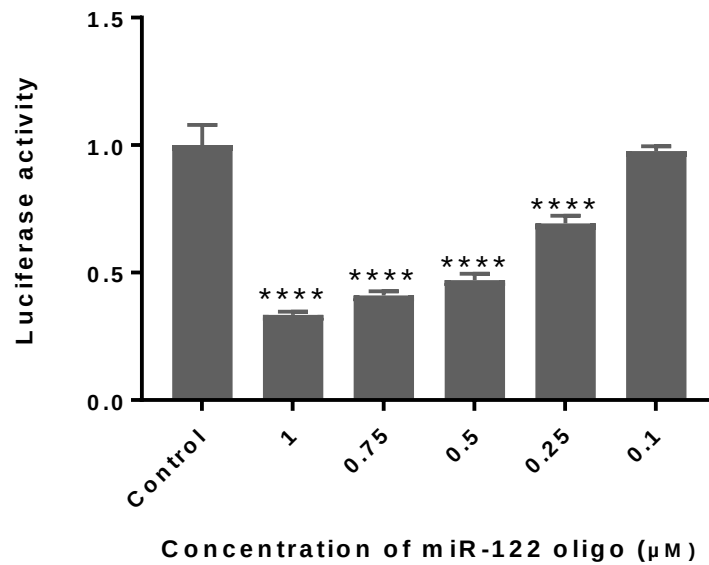


Figure 3.8: Repression of a luciferase reporter with known concentrations of pre-miR-122 oligo. Graphs showing luciferase activity relative to a renilla control. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA compared to pri-miR-122 with *p< 0.05, **p< 0.01, ***p< 0.005, and ****p< 0.0001.

Using the control oligo luciferase assay from figures **3.7** and **3.8**, the concentration of mature miR-122 from different locations of the β -globin vector can be estimated. When expressed from the pri-miR-122 control plasmid, 0.375 μ M of mature miR-122 is produced, with approximately 0.5 μ M of mature miR-122 expressed from the 3' location of the β -globin construct. From the exon 3 location of the β -globin vector approximately 1 μ M of mature miR-122 is expressed. Calculating the expression of mature miR-122 from the intron 2 and exon 2 locations of the β -globin construct is difficult as it falls in the region where the levels of repression plateaus; approximately 1-5 μ M of mature miR-122 is expressed from intron 2, and 5-10 μ M from exon 2. A plateau of repression is reached at approximately 30 % of control luciferase activity, which explains why there is little difference between the different plasmids. Based on this data, it can be concluded that the mature miR-122 expressed from different locations of the β -globin reporter is functional and is able to bind to miR-122 binding sites and repress translation.

3.4: Intronic and exonic miRNA biogenesis and splicing competition

To understand whether miRNA expression affects the levels of spliced and unspliced host mRNA, production of spliced mRNA and unspliced pre-mRNA

from the β -globin construct used in the previous work was analysed by qPCR. Primers spanning the exon-exon junction were used to analyse spliced mRNA, while primers spanning exon 2 and intron 2 were used to analyse unspliced mRNA (**Figure 3.9**). As a control, plasmids containing Drosha-cleavage impaired pre-miR-122 at the same locations were used.

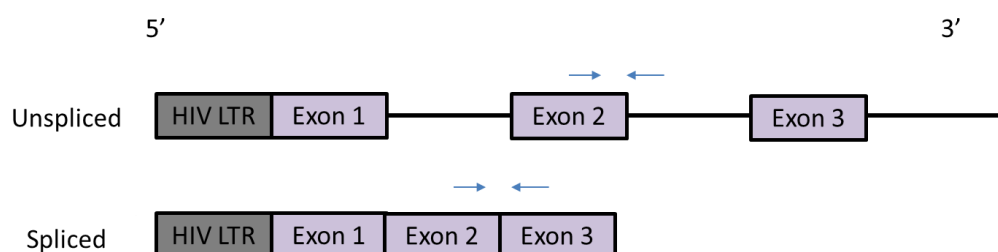


Figure 3.9: Schematic of β -globin mRNA qPCR primer binding sites. The blue arrows indicate the locations where the primers bind.

When pre-miR-122 is inserted within exon 2 or exon 3 of the β -globin construct, the levels of spliced β -globin decreases when compared to the empty vector control; this would be expected if the exon was used as a substrate for Microprocessing (**Figure 3.10**). Similarly, Microprocessor cleavage from an intron would not be expected to have much effect on the β -globin spliced transcript. This data suggests that the levels of spliced host transcript may be affected when miR-122 is expressed from a coding exon, such as exon 2 and exon 3.

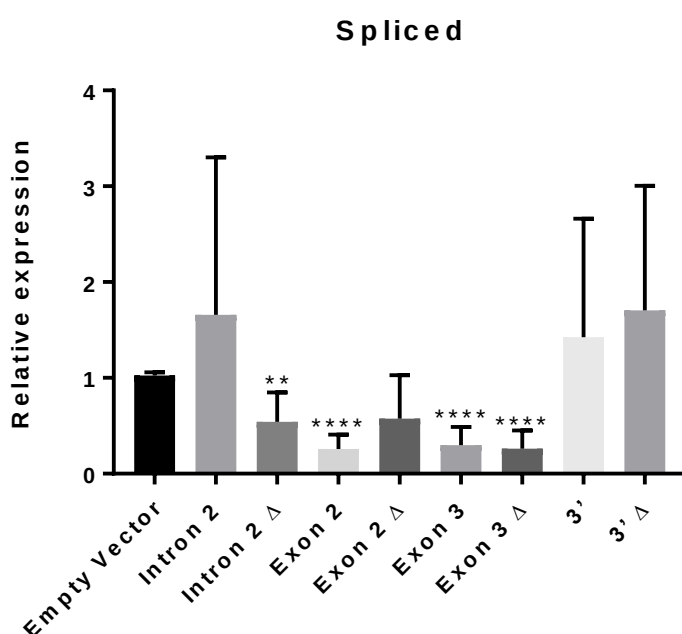


Figure 3.10: Intronic and exonic miR-122 expression from a β -globin plasmid affects host transcript splicing. A graph representing the relative expression of spliced β -globin transcript analysed by qPCR and compared to an

actin control. Δ represents Drosha-cleavage impaired miR-122 Error bars represent standard deviation with a minimum of 5 biological replicates and p-values calculated using a t-test compared to pri-miR-122 with *p< 0.05, **p< 0.01, ***p< 0.005, and ****p, 0.001.

There is a lot of variability within the repeats, especially looking at intron 2 and the 3' plasmids. Also the levels of spliced β -globin decreased when Δ pre-miR-122 is located within an exon which is unexpected. This indicates that any insertions at these locations may affect splicing.

As an alternative method of analysis to look at the levels of spliced β -globin transcript when pre-miR-122 is expressed from different locations, the same RNA samples were analysed by northern blot using a radiolabelled probe that binds to spliced β -globin (**Figure 3.11**). As a negative control, HeLa cells transfected with pri-miR-122 were used and should not express β -globin.

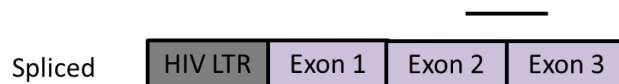


Figure 3.11: A schematic showing the binding location of the northern blot probe to the β -globin transcript. Probe binding location indicated by a black line.

Similarly to the qPCR data, the northern blot analysis shows that the levels of spliced β -globin decreases when pre-miR-122 is expressed from an exon (**Figure 3. 12**). As seen in the qPCR data, the levels of spliced β -globin also decreases when Δ pre-miR-122 is inserted into the plasmid.

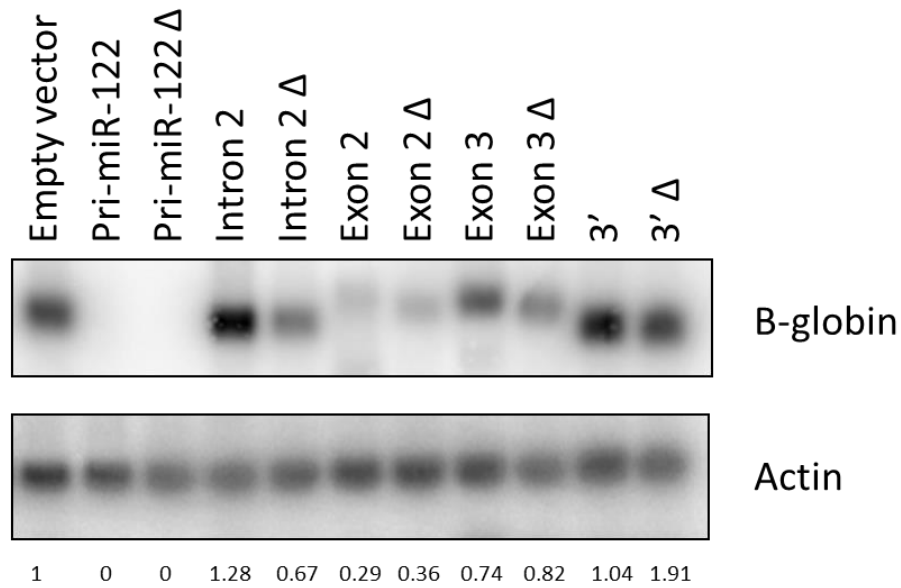


Figure 3.12: The level of spliced β -globin transcripts are affected by the expression of pre-miR-122. Northern blot analysis showing the expression of spliced β -globin transcripts when pre-miR-122 is located with in an intron or an exon compared to an empty β -globin vector. miR-122 is Inc-pri-miR-122. Blot quantification is shown below the blot.

Repeated northern blots also show similar results (**Figure 3.13**), except for when pre-miR-122 is expressed from the 3' location of the β -globin transcript. It is worth noting that figure **3.13A** is missing an intron 2 Δ

sample, and figures **3.13 C** and **D** are missing all the β -globin drosha-cleavage impaired samples.

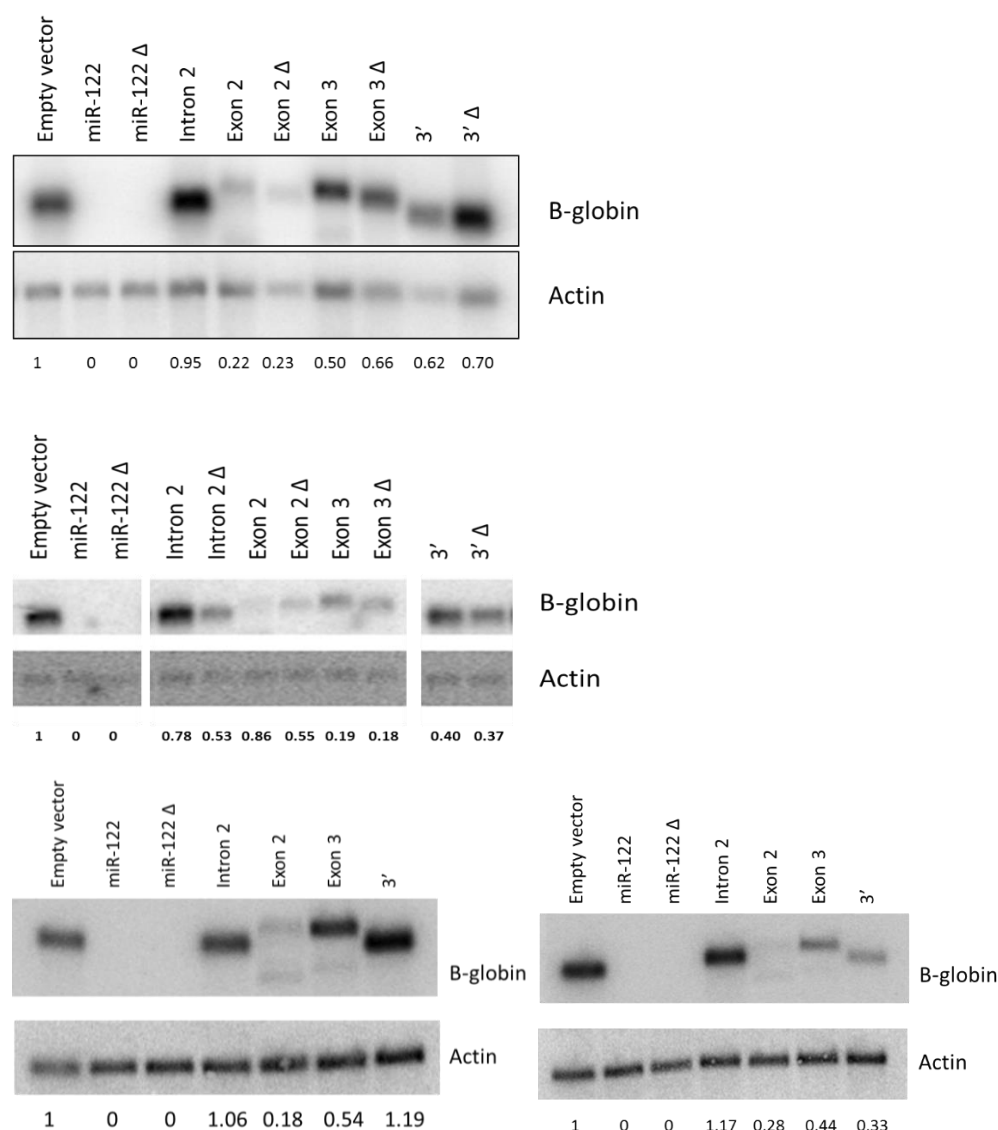


Figure 3.13: Multiple northern blots show that the level of spliced β -globin transcripts is affected by the expression of pre-miR-122. Northern blot analysis showing the expression of spliced β -globin transcripts when pre-miR-122 is located with in an intron or an exon compared to an empty β -globin vector.

The northern blot quantifications were combined as another method of visualising the data (**Figure 3.14**). The data shows a significant decrease in the levels of spliced β -globin transcript when pre-miR-122 is expressed from exon 2 and exon 3. A decrease in the levels of spliced β -globin when Δ pre-miR-122 is expressed from exon 2 compared to the empty vector control indicates that any insertion into this site would have an effect on splicing and therefore it is difficult to draw strong conclusions. As the Drosha cleavage impaired pre-miR-122 inserted within exon 2 has effect

on host splicing, it suggests the insertion itself is having an effect on splicing. The northern blot data largely agrees with the qPCR data shown in figure 3.10 showing that miRNA processing from an exon, but not an intron, decreases production of the spliced host transcript. Interestingly, even though the level of spliced β -globin is reduced compared to the empty vector control, spliced β -globin is still strongly expressed when pre-miR-122 is expressed from an exonic location. This indicated that both mature miR-122 and spliced transcript can be expressed at high levels from the same plasmid when the miRNA is expressed from an exon.

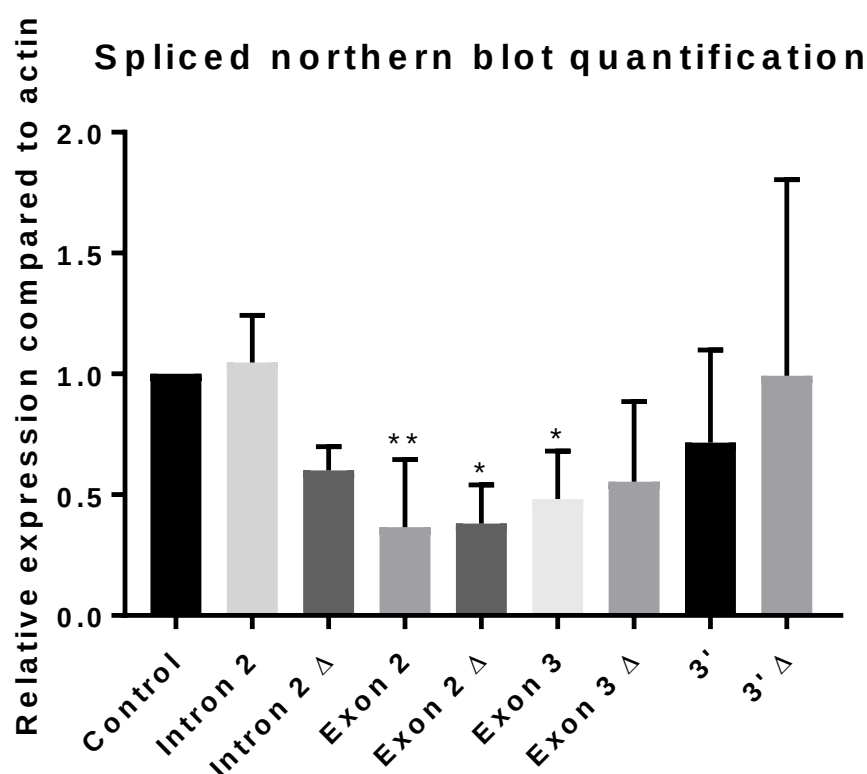


Figure 3.14: Combined northern blot analysis shows that miR-122 expression from an exon leads to a decrease in the levels of host transcript splicing. A graph representing the expression of spliced β -globin transcript analysed and quantified by northern blot. Error bars represent standard deviation with a minimum of 2 biological replicates and p-values calculated using a one-way ANOVA compared to pri-miR-122 with * $p < 0.05$ and ** $p < 0.01$.

Unexpectedly, unspliced levels of β -globin transcript show a increase when pre-miR-122 is expressed from intron 2 (**Figure 3.15**) compared to the empty vector control suggesting that the expression of miR-122 from an intron could be increasing the levels of transcription. Levels of unspliced transcript when pre-miR-122 is expressed from other locations are slightly increased, compared to the empty vector control, and show very similar

levels of expression. A t-test comparing the levels of unspliced β -globin from exon 3 to the empty vector control suggests that this increase is significant. This suggests that there may be an effect on the levels of unspliced host transcript when miR-122 is expressed from an exon, but there is an increase in the levels of unspliced transcript when miR-122 is expressed from an intron suggesting an effect of microprocessing on transcription. However, the variability between replicates is high and so it is difficult to conclude much from this data.

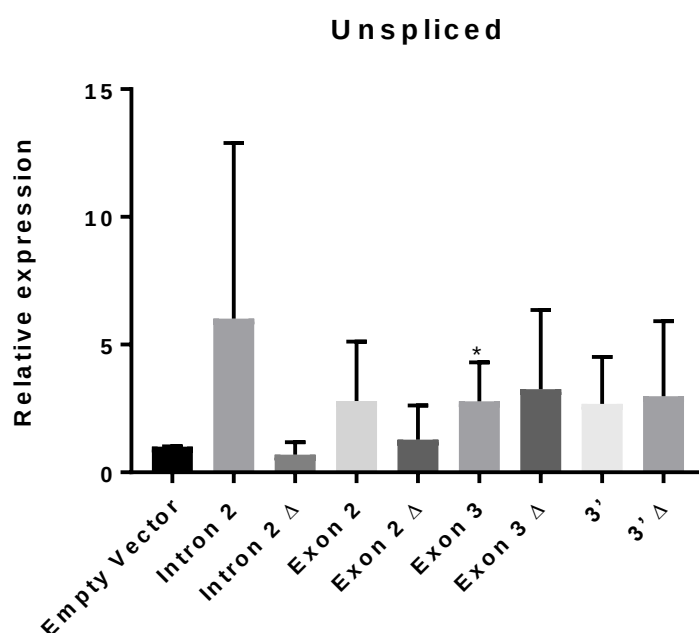


Figure 3.15: Intronic and exonic miR-122 expression may increase the levels of unspliced host transcript. A graph representing the relative expression of unspliced β -globin transcript analysed by qPCR and compared to an actin control. Δ represents Drosha-cleavage impaired miR-122. Error bars represent standard deviation with a minimum of 5 biological replicates and p-values calculated using a t-test compared to pri-miR-122 with * $p < 0.05$.

3.5: Intronic and exonic miRNA biogenesis and Microprocessor knockdown on splicing competition

When looking at the expression of spliced and unspliced β -globin which contain pre-miR-122 there is a lot of variability between repeats as seen by the qPCR and northern blot data. To confirm that the differences between the wild type and Δ pre-miR-122 plasmids are due to Microprocessing, and not the length of the insert, siRNA targeting DGCR8 was transfected into HeLa cells for 72 hours and the pre-miR-122 containing β -globin plasmids transfected 24 hours before harvesting. DGCR8 knockdown has been shown to have a stronger effect on the

Microprocessor compared to Drosha knockdown; Drosha knockdown has been shown to increase the levels of DGCR8 mRNA while DGCR8 knockdown also decreases the level of Drosha protein^{111,119,120}.

Mature miR-122 expression was analysed by qPCR and this data can be analysed in two different ways, either by comparison to the expression of mature miR-122 expressed from the pri-miR-122 control plasmid treated with control siRNA, or, each DGCR8 siRNA treated sample compared to its equivalent siRNA treated control.

The qPCR data shown in figure **3.16** shows the DGCR8 siRNA relative to the control siRNA sample for each pre-miR-122 expressing β -globin plasmid. This data indicates that Microprocessor knockdown efficiently decreases the levels of mature miR-122 expressed when pre-miR-122 is located within the β -globin construct.

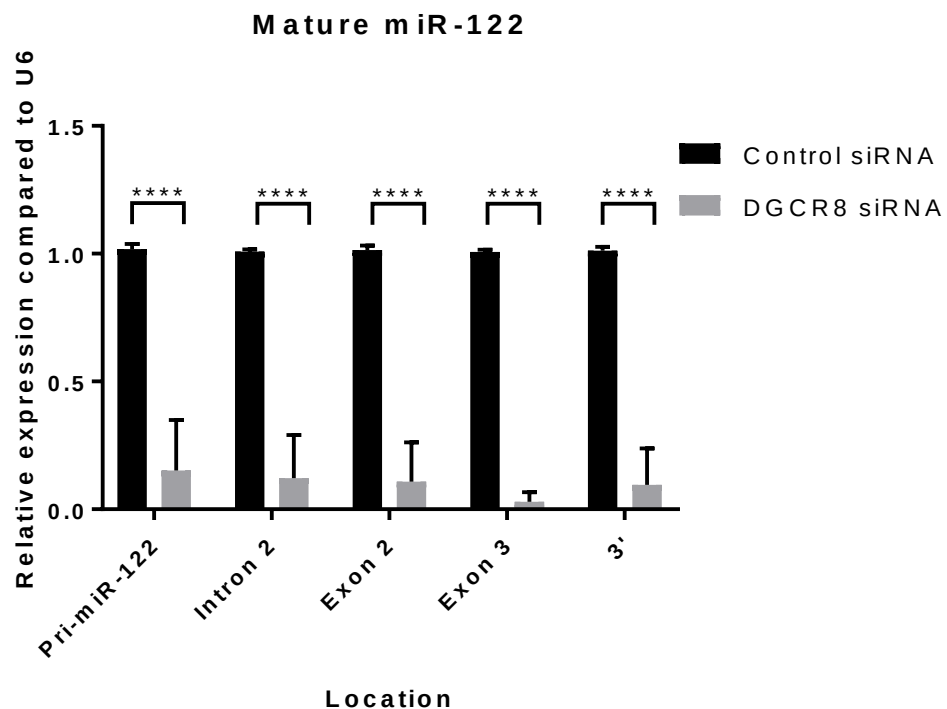


Figure 3.16: Mature miR-122 expression from β -globin reporters containing pre-miR-122 following Microprocessor knockdown. qPCR analysis showing the expression of miR-122 when located with in an intron or an exon compared to endogenous expression under the same reporter following Microprocessor knockdown. All DGCR8 siRNA treated samples are compared to the equivalent control siRNA treated samples. Error bars represent standard deviation with a minimum of 4 biological replicates and p-values calculated using a one-way ANOVA compared to pri-miR-122 with: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

RNA extracted from HeLa cells transfected with a pre-miR-122 β -globin plasmid treated with DGCR8 siRNA to knockdown the Microprocessor was also analysed by northern blot (**Figure 3.17**). Unfortunately, the U6 control exposure was not taken, and as such it is difficult to draw any conclusions from this data as it lacks the loading control, although, equal concentrations of RNA were loaded. This blot indicates that mature miR-122 follows the same trend in mature expression as indicated by qPCR (**Figure 3.16**) and confirms that DGCR8 knockdown was effective in reducing miR-122 expression to undetectable levels.

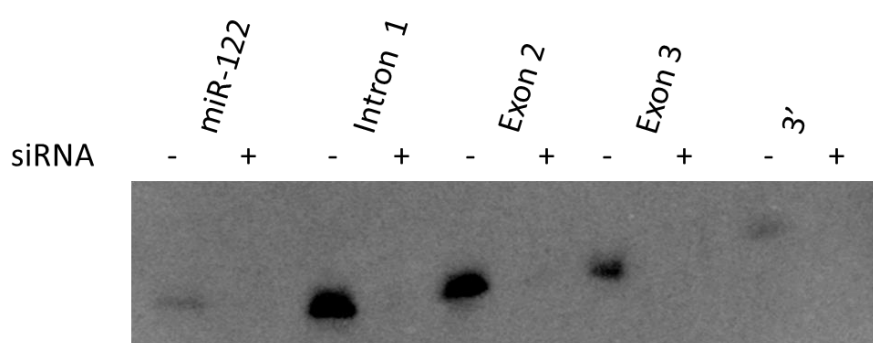


Figure 3.17: Mature miR-122 expression when located within a coding region following Microprocessor knockdown. Northern blot analysis showing the expression of miR-122 when located with in an intron or an exon after Microprocessor knockdown compared to endogenous expression under the same reporter.

Similarly to the qPCR data looking at the expression of mature miR-122 from different locations of a β -globin reporter following siRNA treatment, the levels of spliced and unspliced β -globin transcripts can be analysed either by comparing all the samples to the empty β -globin control treated sample, or by comparing each DGCR8 knockdown sample to its own control siRNA treated sample.

When comparing samples to the control siRNA treated empty vector control, there are no statistically significant differences between any of the samples when pre-miR-122 is expressed from different location of the β -globin plasmid (**Figure 3.18**). However, the levels of control siRNA treated spliced β -globin shows a similar pattern of expression when the same plasmids are expressed from HeLa cells without siRNA transfection (**Figure 3.10**). Microprocessor knocked down HeLa cells expressing pre-miR-122 from intron 2, exon 2 and exon 3 show similar levels of spliced mRNA

compared to the control, while 3'-pre-miR-122 shows an increased amount of spliced mRNA. The significance of these differences may be lost due to the variability in the other samples. More repeats may help improve the significance of this data, however, it still supports the idea that when miRNA is located within an exon, both mature miRNA and spliced mRNA can be produced.

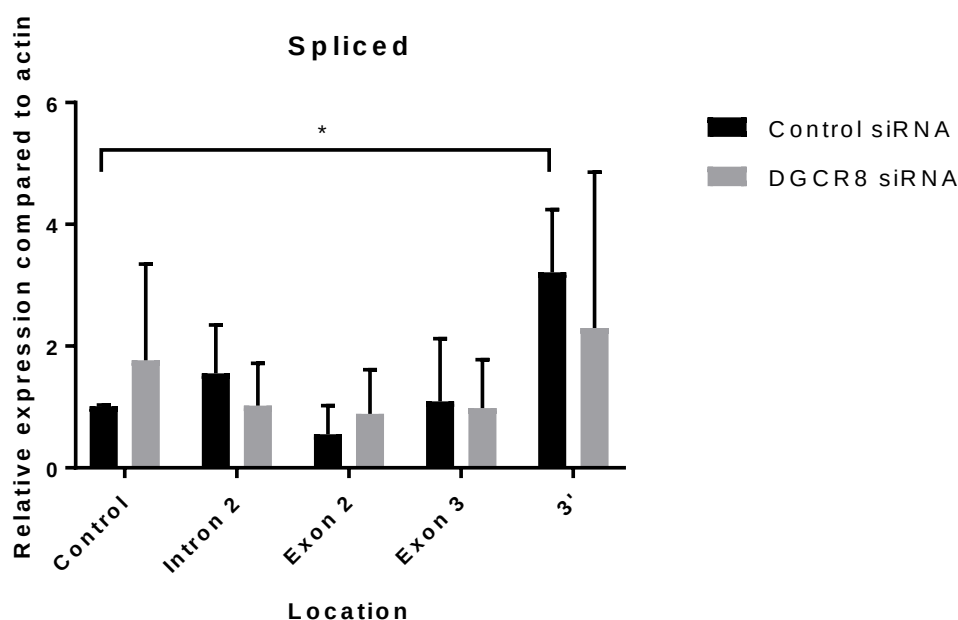


Figure 3.18: Intronic and exonic miR-122 expression may affect host transcript splicing following Microprocessor knockdown. A graph representing the relative expression of spliced β -globin transcript analysed by qPCR and compared to an actin control. Each sample is compared to the control siRNA treated control. Error bars represent standard deviation with a minimum of 5 biological replicates and p-values calculated using a one-way ANOVA compared to vector control siRNA with * $p < 0.05$.

To further analyse this data, each DGCR8 knockdown sample was compared relative to its own control siRNA sample; the siRNA control has been set to 1 and is indicated by the dotted line (**Figure 3.19**). When the levels of spliced β -globin are compared to its own control siRNA, for example, DGCR8 siRNA treated exon 2-pre-miR-122 is compared to the control siRNA treated exon 2-pre-miR-122 sample, the data indicates that expression of pre-miR-122 from the 3' flanking region has little or no effect on the levels of spliced β -globin. When pre-miR-122 is located within intron 2 or exon 3, the levels of spliced mRNA decrease following Microprocessor knockdown, whereas an increase is shown for exon 2 and exon 3. However, these differences are not significant, and as previously

discussed, this may be due to the variation in repeats and may be improved by repeating the experiment.

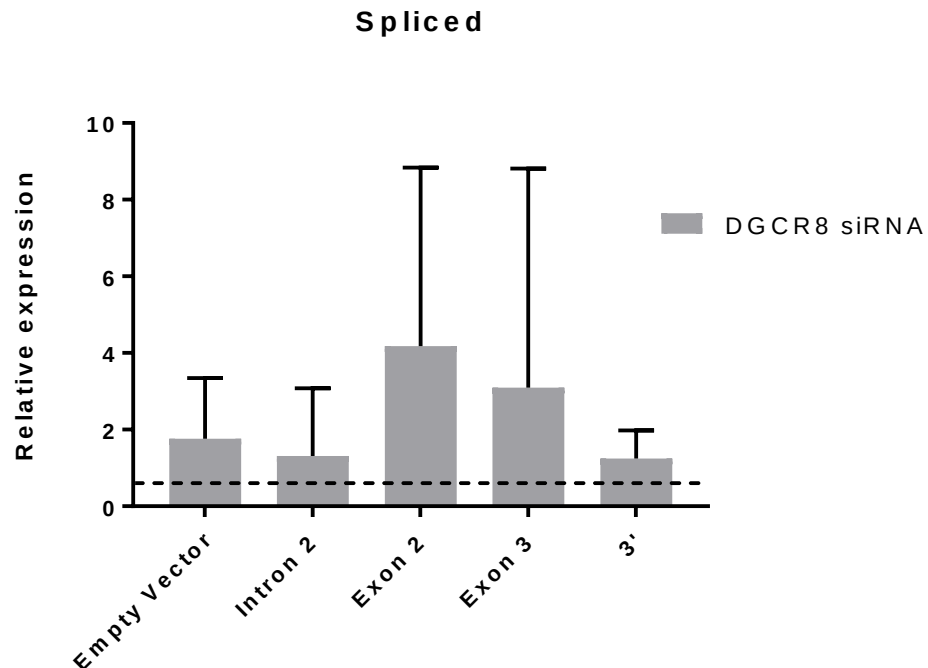


Figure 3.19: MiR-122 expression from an exon or intron may affect host transcript splicing following Microprocessor knockdown. A graph representing the relative expression of spliced β -globin transcript analysed by qPCR and compared to an actin control. All samples are compared to their control siRNA treated equivalent which are set a 1 and are represented by the dashed line. Error bars represent standard deviation with a minimum of 6 biological replicates and p-values calculated using a one-way ANOVA compared to vector siRNA control.

To confirm this data, the same samples were analysed by northern blot (**Figure 3.20**). Each northern blot was quantified to aid analysis, however, due to the background in the blots there are a lot of variation between the three blots.

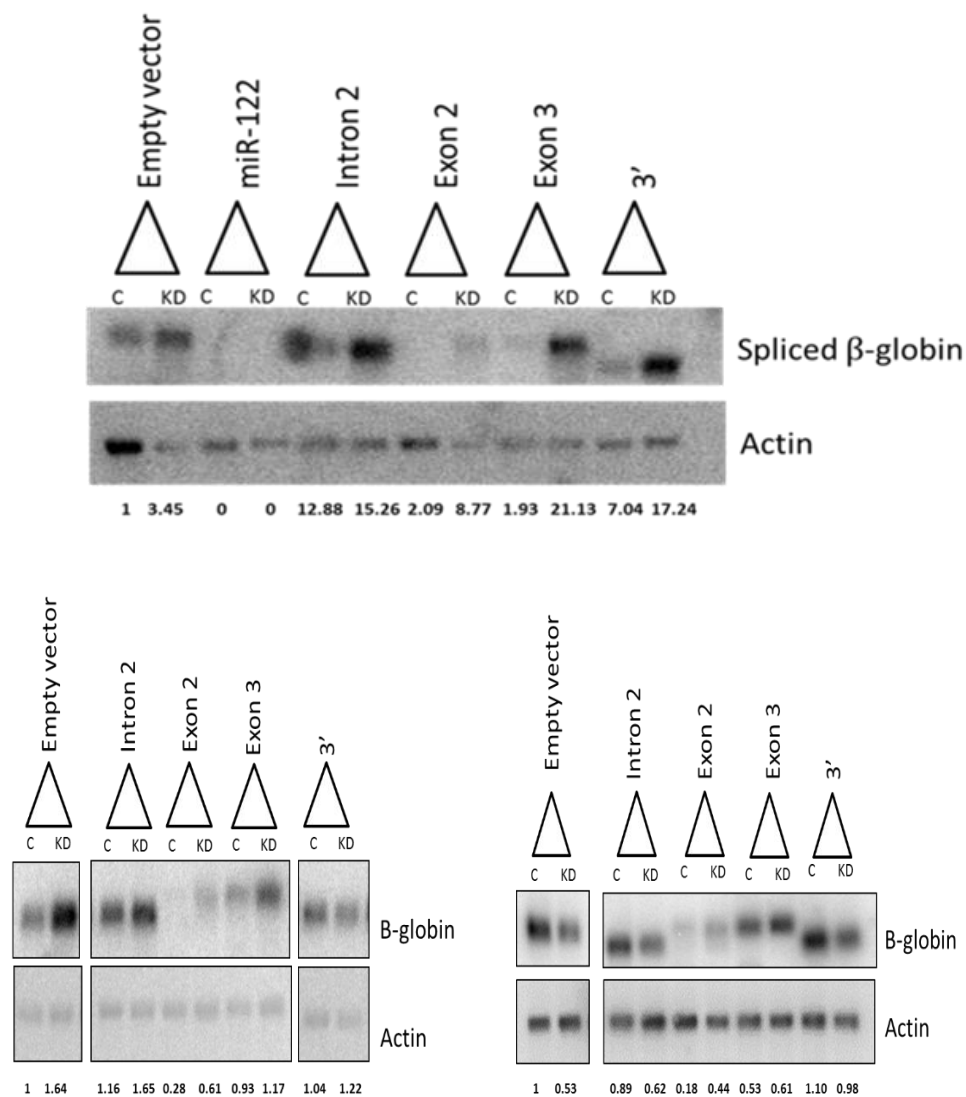


Figure 3.20: The level of spliced β -globin transcripts are affected by the expression of pre-miR-122 following Microprocessor knockdown. A) and B) and C) Northern blot analysis showing the expression of spliced β -globin transcripts when pre-miR-122 is located within an intron or an exon compared to an empty β -globin vector following DGCR8 siRNA treatment.

The northern blot quantifications were collated and shown as a graph to aid interpretation (**Figure 3.21**). While the graph shows that there is a lot of variation within some samples, some samples have relatively little error. As with the qPCR data shown in figure **3.10**, the levels of spliced β -globin control treated siRNA samples follow the same trend; an increase of spliced β -globin transcript is observed when intron 2 contains pre-miR-122 compared to the empty vector control, followed by a decrease in transcript when pre-miR-122 is expressed from exon 2 and exon 3 compared to the

empty vector control. Due to the variation within the DGCR8 siRNA treated samples, it is difficult to draw any conclusions.

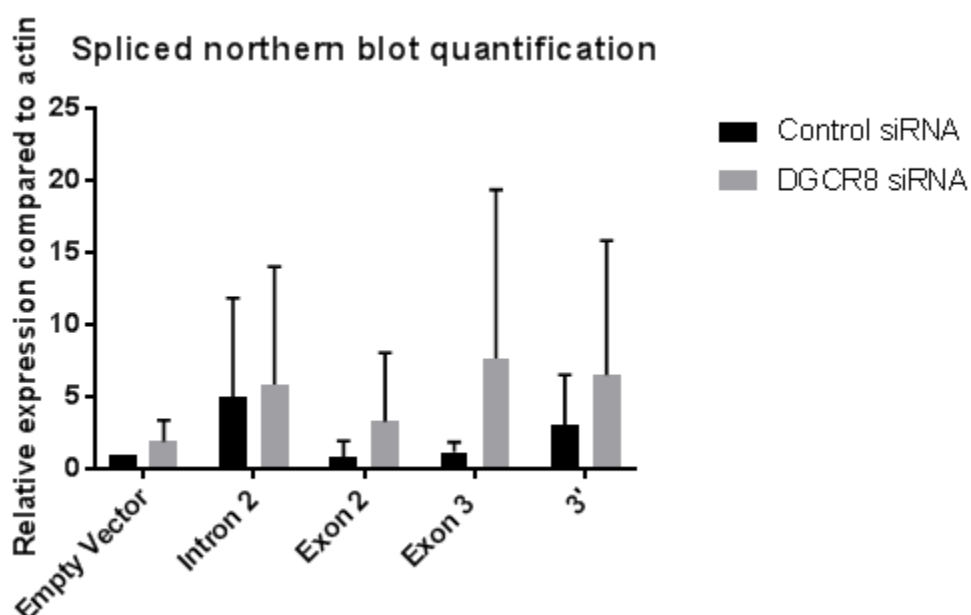


Figure 3.21: A graph visualising the northern blot quantification of spliced β -globin. Data analysed is from the northern blots shown in figure 3.22. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA compared to pri-miR-122.

The levels of unspliced β -globin following siRNA treatment were also analysed by qPCR. The control siRNA treated samples show an increase in the levels of unspliced β -globin when pre-miR-122 is located within intron 2 (**Figure 3.22**), although this is a smaller increase compared to the data shown in figure 3.15. However, the effect of expressing pre-miR-122 from different locations within the β -globin transcript may be masked by the variation between data sets. Similarly, there is a lot of error within the samples treated with DGCR8 siRNA to knockdown the Microprocessor and it is difficult to form any conclusions about this data.

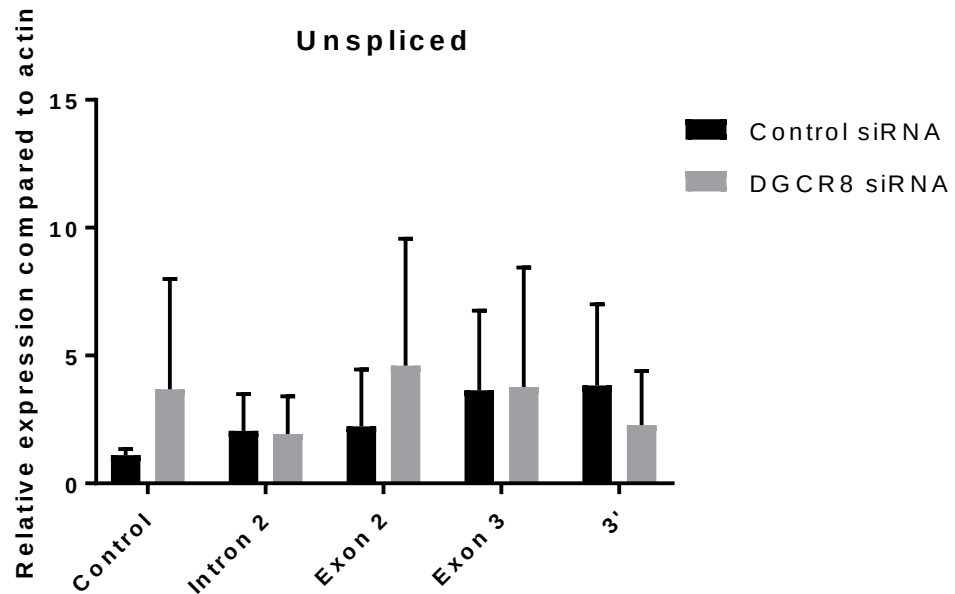


Figure 3.22: Intronic and exonic miR-122 expression may affect host transcript splicing following Microprocessor knockdown. A graph representing the relative expression of unspliced β -globin transcript analysed by qPCR and compared to an actin control. Each sample is compared to the control siRNA treated control. Error bars represent standard deviation with a minimum of 4 biological replicates and p-values calculated using a one-way ANOVA compared vector siRNA control.

To try to elucidate whether there is an effect on the levels of unspliced β -globin mRNA when pre-miR-122 is located within the transcript, DGCR8 knockdown samples were compared relative to the control treated siRNA samples and represented by a dotted line (**Figure 3.23**). This data suggests that the presence of the pre-miR-122 hairpin, and also Microprocessing may have an effect on the levels of unspliced transcript.

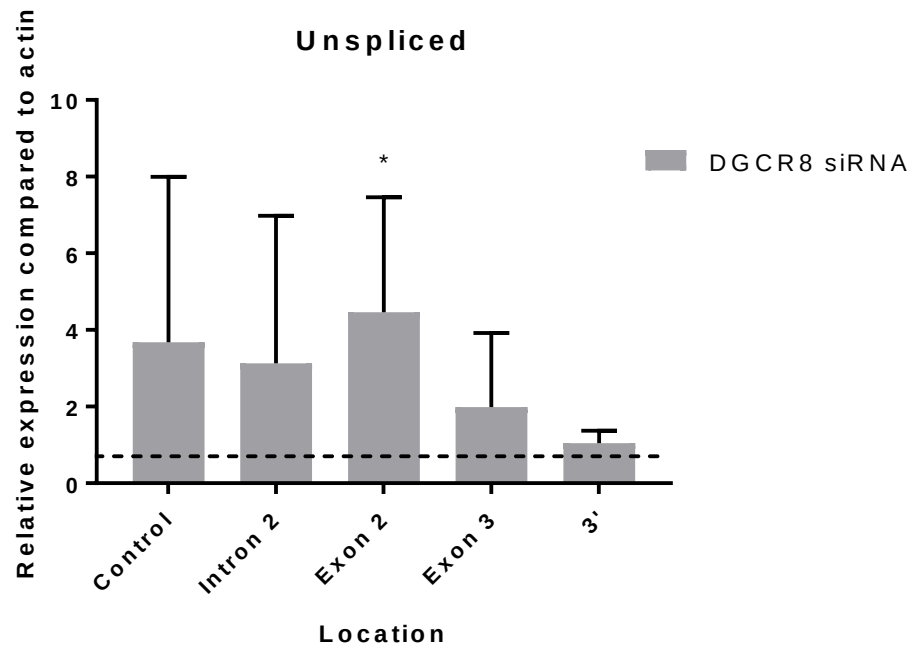


Figure 3.23: Intronic and exonic miR-122 expression may affect host transcript splicing following Microprocessor knockdown when compared to DMSO treated cells. A graph representing the relative expression of spliced β -globin transcript analysed by qPCR and compared to an actin control. Each sample is compared to its own control siRNA treated sample which is set to 1 and is represented by the dashed line. Error bars represent standard deviation with a minimum of 4 biological replicates and p-values calculated using a one-way ANOVA compared vector siRNA control.

3.6: Drosha cleavage efficiency

As another measure of looking at how pre-miR-122 is processed when it is located at different locations within the β -globin vector, qPCR was used to measure the drosha cleavage efficiency. qPCR primers were designed to span across the 5' drosha cleavage site for intron 1, exon 2 and 3' flanking region β -globin primers, with one primer located within the miR-122 hairpin sequence. This will measure how much miR-122 hairpin is retained within the mRNA transcript, with a decrease in qPCR signal indicating that more Microprocessing has occurred (**Figure 3.24**).

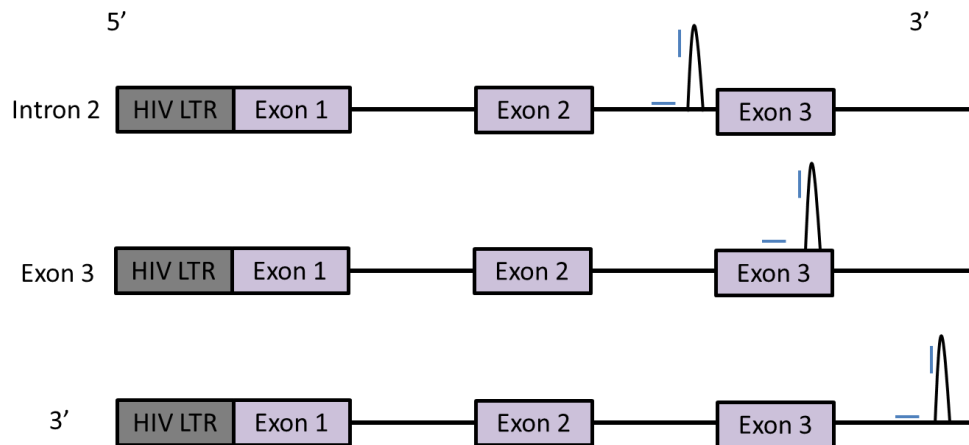


Figure 3.24: Schematic of qPCR primer bind sites used to measure Drosha cleavage efficiency. Primer binding sites indicated by dark blue lines.

HeLa cells were treated with DGCR8 siRNA before transfecting with one of the above mentioned plasmids. The qPCR data is analysed with the control siRNA samples being compared to the equivalent DGCR8 siRNA samples, and the DGCR8 knockdown samples set at 1 as this should be where hairpin retention is at its highest (**Figure 3.25**). The data suggests that drosha cleavage is more efficient when the miR-122 hairpin is located within intron 2. This could indicate that miRNA processing is more efficient from an intron compared to other locations. It could also indicate that a hairpin located earlier in the transcript is available for Microprocessing for longer compared to hairpins located downstream of the transcript.

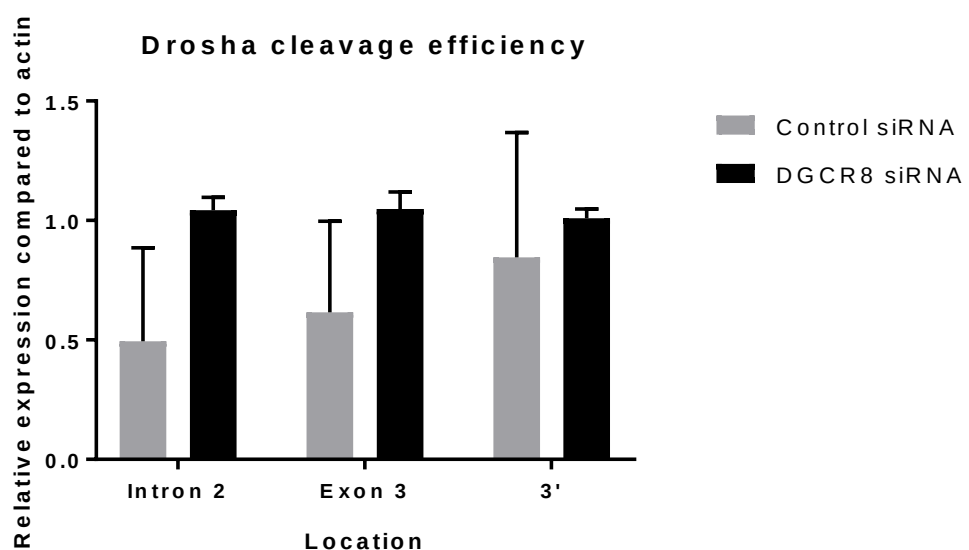


Figure 3.25: DGCR8 knockdown affects Drosha cleavage efficiency. A graph representing the relative expression of miR-122 hairpin retention analysed by qPCR

and compared to an actin control. Each sample is compared to its own control siRNA treated sample which is set to 1 and is represented by the dashed line. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a t-test.

3.7: Discussion

This chapter looks at the effect expressing pre-miR-122 from intron 2, exon 2, exon 3 or the 3' flanking region of a β -globin construct has on the host transcript and mature miR-122 expression with the aim of providing a direct comparison of how the same miRNA is expressed from different genomic contexts. To understand the effects of miRNA location on miRNA processing and splicing of these transcripts and of miRNA function, qPCR, northern blot and luciferase assays were used.

The results show that mature miR-122 expression is increased when expressed from a coding gene; expression from an intron or exon is increased compared to the lncRNA as well as the 3' location. This was also confirmed using a luciferase assay which also showed that the location from which a miRNA is expressed has no effect on its downstream function.

While mature miR expression from an intron was expected, it is surprising that mature miR-122 is expressed similarly from an exon. High expression of mature miR-122 from an exon would suggest that the host β -globin transcript would be destroyed in the process, with the host transcript then being targeted for degradation. MiRNA expression from an exon could also lead to exon skipping. The northern blots suggests that the spliced β -globin transcript detected is full length, however a smaller, and weaker, band can be detected on some of the northern blots when pre-miR-122 is expressed from an exon. This suggests exon skipping does occur.

Due to the large amount of error in the DGCR8 knockdown experiments, it is difficult to draw conclusions from this data. This error may have been caused from DNA contamination, or siRNA and plasmid transfection occurring at the same time may be having an effect on the cells, such as cell death.

Similar levels of mature miR-122 have been shown to be produced from exon 2 and exon 3. However, following Microprocessor knockdown, the levels of spliced β -globin increase compared to the control knockdown

when pre-miR-122 is located within exon 2, but not when pre-miR-122 is located within exon 3. An increase in the levels of spliced transcript when the Microprocessor is knocked down could indicate that the miRNA hairpin is being processed prior to splicing, which has been described in the literature looking at intronic miR biogenesis^{198,237}. This would suggest that there is competition between splicing and miRNA biogenesis from an exon, such as exon2 of the β -globin gene, and that the levels of transcription may be increased to allow for both products to be generated.

As a control, pre-miR-122 was inserted into the 3' flanking region of the β -globin vector. While this site is 276 bp downstream of the PAS, this is the region where RNA polymerase II would be terminating transcription. Interestingly, mature miR-122 is produced from the 3' flanking region is at a similar level to the lncRNA control. In both these circumstances, the miRNA hairpin is at the 3' end of the transcript and it is possible that distance along the transcript may influence miRNA biogenesis.

So far, the data suggests that the location of a miR-122 hairpin in the β -globin transcript can affect the expression of the mature miR-122, with more mature miR-122 expressed from an intron or exon, compared to the 3' flanking regions or the lncRNA control. The data also suggests that hairpins located towards the 5' end of the transcript are more efficiently processed, with less miRNA being processed from locations towards the end of a transcript. As Microprocessing is known to occur before splicing, this may also affect the efficiency of splicing on the mRNA. While the data shown in this chapter shows some evidence for competition between miRNA biogenesis and splicing, it also unexpectedly shows that both the spliced mRNA and mature miRNA can be effectively produced from the same plasmid when a miRNA is in an exon. More work is needed, to understand how miRNA processing from an intron or exon may affect splicing efficiency.

Rather than looking at total RNA, looking at newly transcribed RNA using 4-SU labelling or chromatin-associated RNA may provide a more informative way of measuring the effect of miRNA processing on β -globin containing pre-miR-122. This may also help reduce the variability that has been seen in the unspliced and spliced data. However, preliminary work looking at chromatin-associated RNA from HeLa cells transfected with a

plasmid proved difficult as highly levels of DNA contamination were detected.

Additionally, it would be interesting to look more directly at how the pre-miR-122 position in the β -globin transcript can affect its expression, as hairpins towards the 3' end of the transcript have less time to be processed than a hairpin at the 5' end. To do this, pre-miR-122 could be expressed from different locations of the β -globin gene, such as intron 1 and exon 1. The expression of mature miR-122, or spliced and unspliced β -globin could be compared to the expression from intron 1. Similarly, different sites of insertion within a location could be used, for example, pre-miR-122 could be expressed from the beginning, middle and end of an intron. A different expression system to the β -globin gene could also be used, such as luciferase gene; this would allow miRNA expression to be analysed from a 5' to 3' direction without the influence of splicing. Additionally, a range of qPCR primers could be designed spanning the β -globin transcript, which could be used to calculate how quickly an intron is spliced relative to a miRNA hairpin, and also to assess exon skipping. This could provide information about how position of a miRNA hairpin influences both splicing and Microprocessor cleavage in a temporal manner.



Figure 4.2: Schematic of pri-miR-122

To understand what role splicing may have on the biogenesis of miR-122, Huh7 cells which endogenously express miR-122 can be treated with splice inhibitor drugs such as pladienolide B (Pla B). Similarly, a plasmid containing pri-miR-122 can be transfected into HeLa cells prior to treatment with pla B. By looking at the effect of splice inhibition of the levels of spliced and unspliced pri-miR-122, the relationship between splicing and miRNA biogenesis may be elucidated.

4.2: Pla B optimisation in HeLa and Huh7 cells

Previously published work using pla B as a splicing inhibitor treated cells for 4 hours with a concentration of $1 \mu\text{M}$ ⁷². To ensure that this would be efficient to inhibit splicing in HeLa cells, HeLa cells were first transfected with the pri-miR-122 plasmid used in chapter 3. 24 hours post-transfection, HeLa cells were treated with different concentrations of Pla B, ranging from $10 \mu\text{M}$ to 1 nM for 4 hours. As Huh7 cells endogenously express miR-122, the cells were treated with different concentrations of pla B 24 hours after seeding. After the drug treatment, total RNA was harvested, reversed transcribed and the levels of spliced and unspliced pri-miR-122 analysed by qPCR. The primer binding sites are indicated in figure 4.3. It is worth noting that the unspliced primers amplify within intron 1 and not across the intron-exon boundary and it is possible that excised intron 1 may be detected.

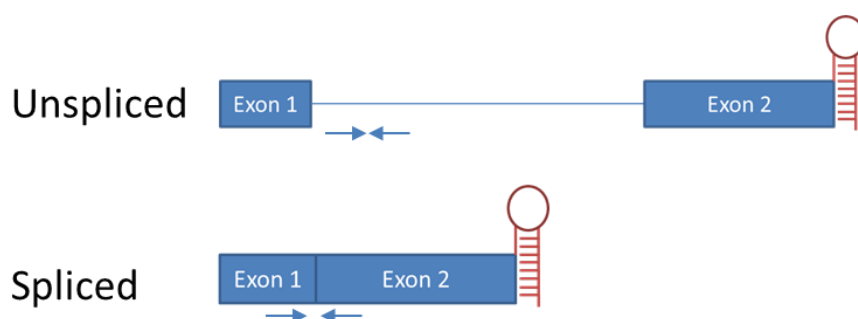


Figure 4.3: A schematic of the qPCR binding sites in pri-miR-122. The blue arrow indicate the regions qPCR primers bind to.

When treated with pla B to inhibit splicing, the levels of spliced pri-miR-122, both endogenous and overexpressed, decreases (**Figure 4.4**). At all concentrations tested, the levels of spliced endogenous pri-miR-122 from Huh7 cells significantly decreased (**Figure 4.4B**). The levels of overexpressed pri-miR-122 from HeLa cells also reduce with pla B, although due to the variability within the samples at 1, 10 and 100 nM the decrease seen at 1 and 10 μ M are not statistically significant (**Figure 4.4A**). This data suggests that pla B inhibits the splicing of pri-miR-122 in Huh7 cells and HeLa cells expressing pri-miR-122.

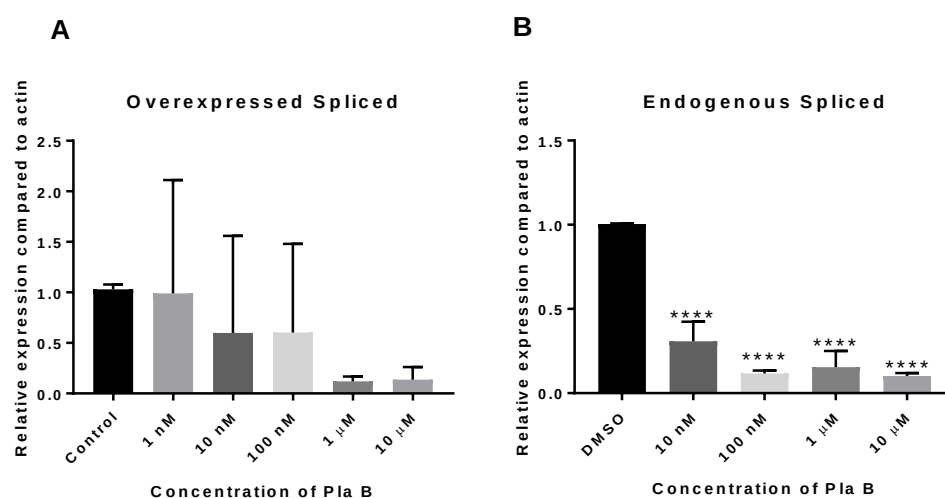


Figure 4.4: Splice inhibition using pla B reduces the levels of spliced pri-miR-122. **A)** A graph representing the relative expression of spliced overexpressed pri-miR-122 transcript from HeLa cells analysed by qPCR and compared to an actin control. **B)** A graph representing the relative expression of spliced endogenous pri-miR-122 transcript from Huh7 cells analysed by qPCR and compared to an actin control. Error bars represent standard deviation with a minimum of 3 biological replicates and significance was measured with a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

Interestingly, qPCR analysis following pla B treatment shows an increase in the levels of unspliced overexpressed pri-miR-122 in HeLa cells, but the opposite effect for endogenous pri-miR-122 in Huh7 cells (**Figure 4.5**). An increase in the levels of unspliced transcript is expected, as less unspliced transcript being processed by splicing should lead to a build-up of unspliced transcript. This effect is seen when pri-miR-122 is overexpressed in HeLa cells (**Figure 4.5A**). Conversely, there is a distinct lack of unspliced pri-miR-122 in Huh7 cells which suggests that splice inhibition by pla B may be having an effect on transcription (**Figure 4.5B**).

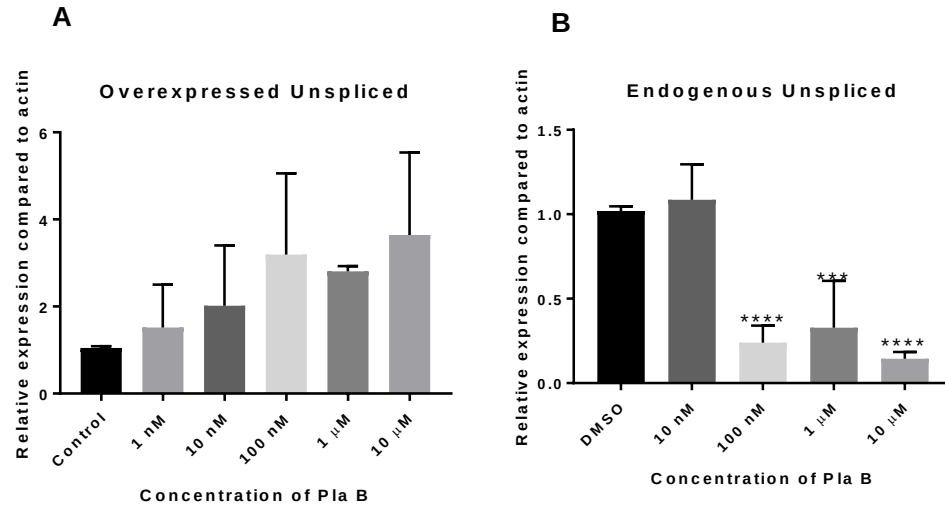


Figure 4.5: Splice inhibition using pla B reduces the levels of endogenous but not overexpressed unspliced pri-miR-122. **A)** A graph representing the relative expression of spliced overexpressed pri-miR-122 transcript from HeLa cells analysed by qPCR and compared to an actin control. **B)** A graph representing the relative expression of spliced endogenous pri-miR-122 transcript from Huh7 cells analysed by qPCR and compared to an actin control. Error bars represent standard deviation with a minimum of 3 biological replicates and significance was measured with a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

From this data, 1 μM of pla B was determined to be an appropriate concentration of drug to use to inhibit splicing for both endogenous and overexpressed pri-miR-122. To further optimise the pla B treatment, the most appropriate length of treatment was determined using qPCR.

When pri-miR-122 is overexpressed in HeLa cells, the levels of spliced pri-miR-122 decreases time wise, with the greatest reduction in spliced transcripts occurring at 8 hours (**Figure 4.6A**). For endogenous pri-miR-122, the levels of spliced pri-miR-122 is reduced to less than 10 % at 2 hr, and little deviation from this is seen for longer treatments (**Figure 4.6B**). Interestingly, a significant reduction in splice pri-miR-122 is seen within 1 hour of treatment with pla B for both cell lines, indicating that inhibition of splicing is quickly apparent on total levels of pri-miR-122.

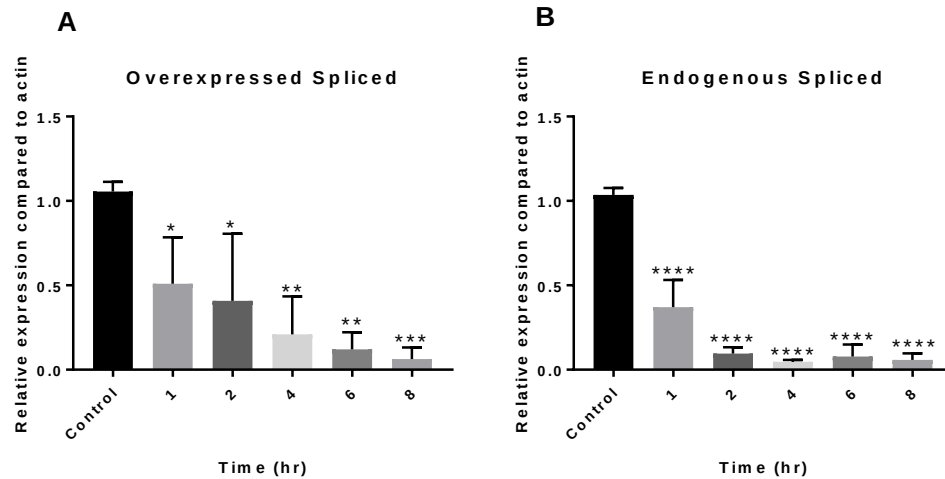


Figure 4.6: Splice inhibition using pla B reduces the levels of spliced pri-miR-122 within 1 hour. A) A graph representing the relative expression of spliced overexpressed pri-miR-122 transcript from HeLa cells analysed by qPCR and compared to an actin control. **B)** A graph representing the relative expression of spliced endogenous pri-miR-122 transcript from Huh7 cells analysed by qPCR and compared to an actin control. Error bars represent standard deviation with a minimum of 3 biological replicates and significance was measured with a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

As seen in figure 4.5, the effect of splice inhibition on unspliced pri-miR-122 is different for the endogenous compared to the overexpressed (Figure 4.7). When HeLa cells overexpressing pri-miR-122 are treated with pla B, the levels of unspliced pri-miR-122 increases and may accumulate over time (Figure 4.7A). However, endogenous pri-miR-122 from Huh7 cells treated with pla B show significant decreases in unspliced transcripts (Figure 4.7B), and similarly to figure 4.6 little difference is seen between 2 and 8 hours.

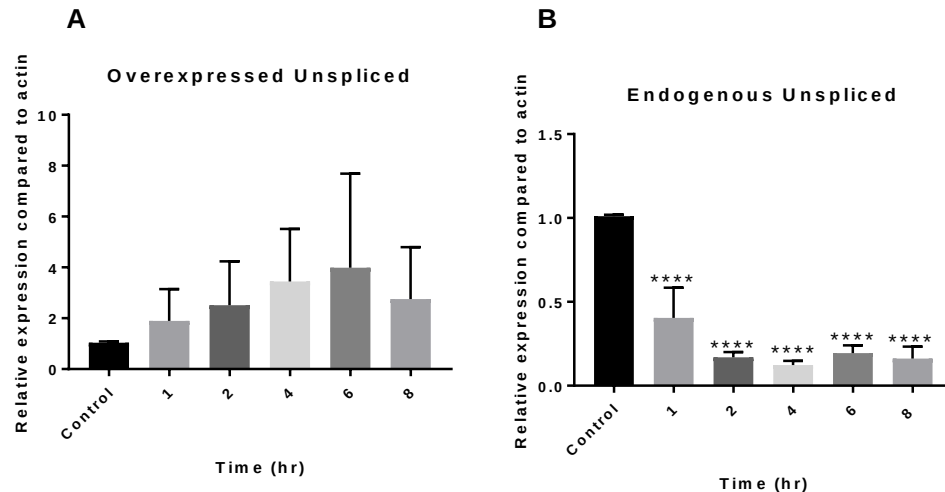


Figure 4.7: Splice inhibition leads to an increase in unspliced pri-miR-122 when overexpressed in HeLa cells. A) A graph representing the relative expression of unspliced overexpressed pri-miR-122 transcript from HeLa cells analysed by qPCR and compared to an actin control. **B)** A graph representing the relative expression of unspliced endogenous pri-miR-122 transcript from Huh7 cells analysed by qPCR and compared to an actin control. Error bars represent standard deviation with a minimum of 3 biological replicates and significance was measured with a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

Based on these data, a treatment of 1 μ M Pla B for 4 hr will be used to inhibit splicing in Huh7 cells, as well as HeLa cells transfected with pri-miR-122 as this will achieve a strong effect on splicing and minimise the possibility of secondary effects.

The levels of mature miR-122 following 4 hours of 1 μ M pla B treatment was also analysed by qPCR. After 4 repeats, the data initially suggests that the level of mature miR-122 may increase within this time both for endogenous mature miR-122 (**Figure 4.8B**) and mature miR-122 expressed from a plasmid in HeLa cells (**Figure 4.8A**). An increase in mature miR-122 may indicate competition between miRNA biogenesis and splicing, with splicing inhibition leading to an increase in miRNA biogenesis. However, with the levels of unspliced and spliced endogenous pri-miR-122 decreasing with pla B treatment, it would seem likely that the levels of mature miR-122 would also decrease as there is less substrate for the Microprocessor to cleave. With the miR-122 known to be very stable and have a long half life, this would be difficult to detect within 4 hours.

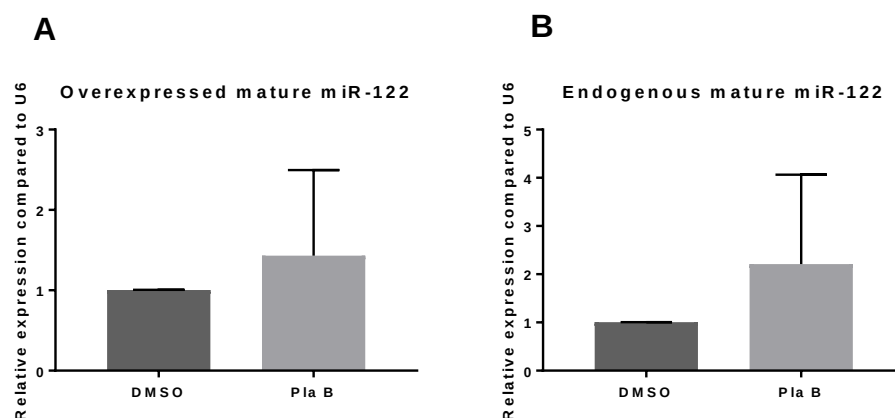


Figure 4.8: Mature miR-122 after splice inhibition. A) A graph representing the relative expression of mature miR-122 from overexpressed pri-miR-122 in HeLa cells analysed by qPCR and compared to a U6 control. **B)** A graph representing the relative expression of endogenous mature miR-122 from Huh7 cells analysed by qPCR and compared to a U6 control. Error bars represent standard deviation with a minimum of 4 biological replicates and significance was measured with a T-test.

As a global inhibitor of splicing, it is necessary to determine whether Pla B has a cytotoxic effect on cell growth. To assess the effect of Pla B on HeLa cells, cells were treated with Pla B and the assay reagent AlamarBlue for either 1, 4 or 24 hours. Resazurin, the active compound in AlamarBlue, is a cell permeable compound which is reduced in metabolically functional cells to resorufin. While resazurin is blue in colour and does not fluoresce, resorufin is a red fluorescent molecule which can be easily detected using a plate reader. Less metabolically active cells will produce less resorufin, and should pla B be cytotoxic to cells, the levels of fluorescence will decrease.

As shown in figure 4.9, it can be determined that HeLa cells incubated with Pla B for 1 or 4 hours show no evidence of cytotoxicity. However, when HeLa cells are treated with pla B for 24 hours there is a small but statistically significant negative effect on cell metabolism, indicating that pla B is cytotoxic to HeLa cells with prolonged treatment. A lack of cytotoxicity after 4 hr of treatment indicates that this is a suitable treatment.

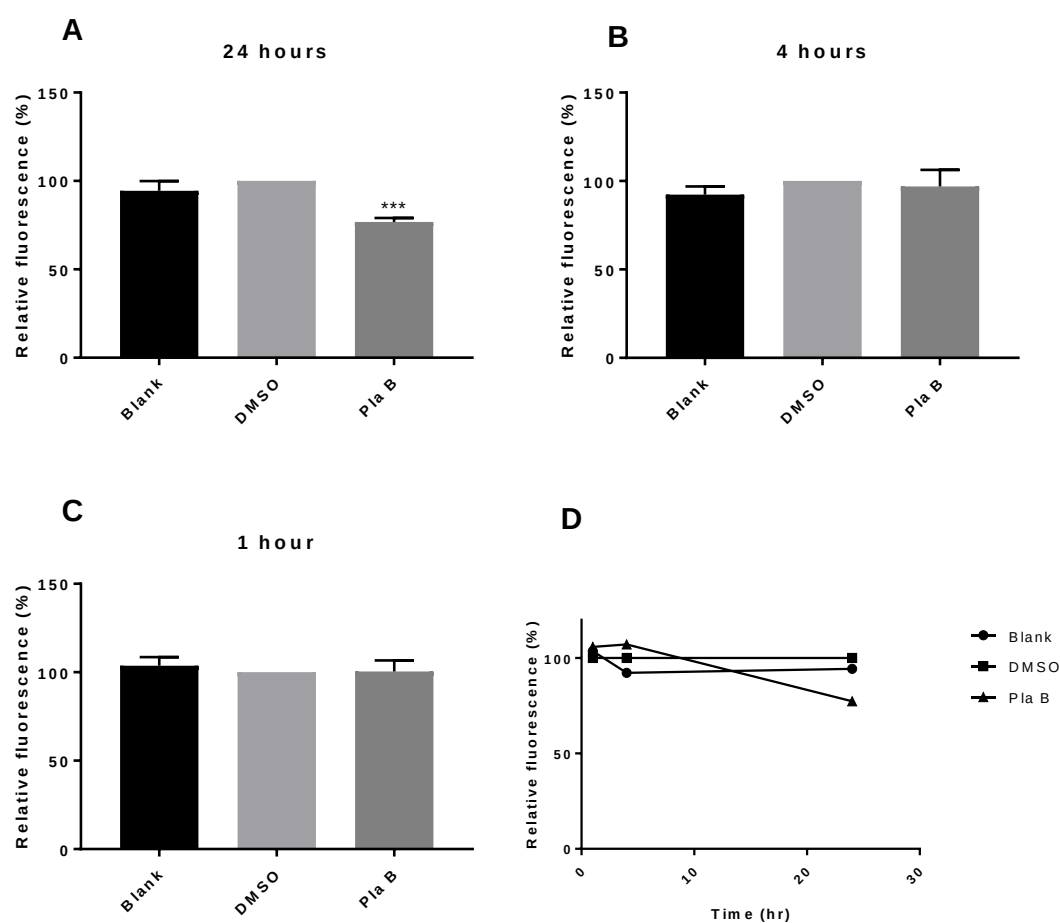


Figure 4.9: Short treatments of pladienolide B are not cytotoxic to HeLa cells. A-C) Graphs showing the relative fluorescence of resorufin of pla B compared to a DMSO control for 24, 4 and 1 hour respectively. **D)** A line graph representing the change in resorufin fluorescence over time. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA compared to pri-miR-122 with * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.005$.

Similarly, the effect of pla B on the viability of Huh7 cells was measured using AlamarBlue (**Figure 4.10**). This data indicates that pla B has little

effect on cell viability even at 24 hours, and also indicates that a treatment of 1 μ M pla B for 4 hours is a suitable to inhibit splicing in Huh7 cells.

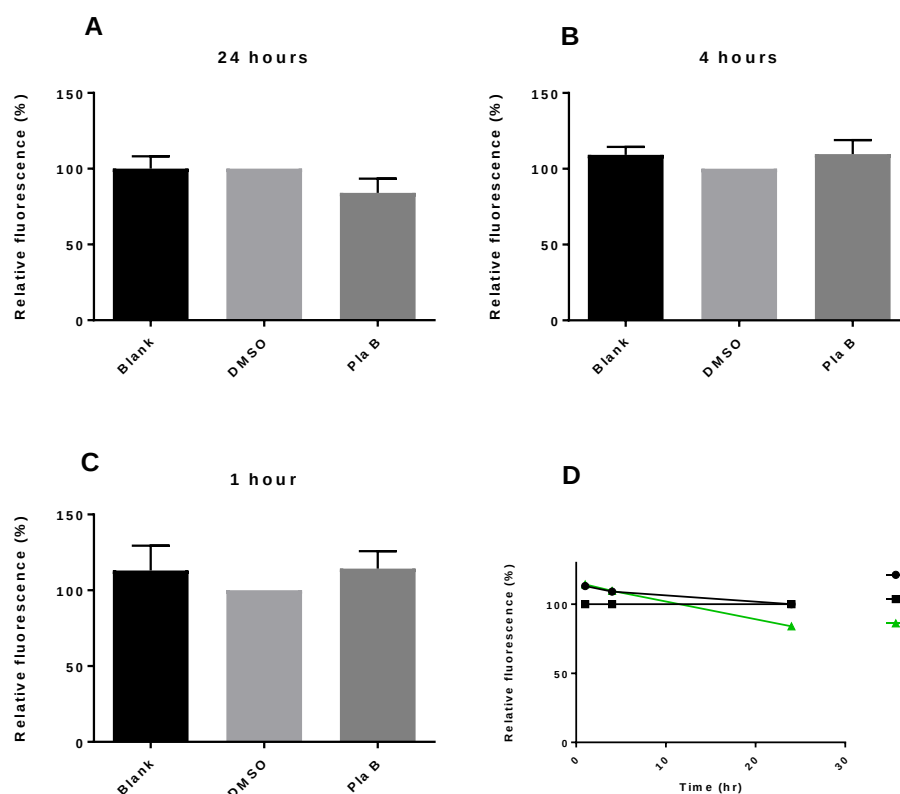


Figure 4.10: Short treatments of pladienolide B are not cytotoxic to Huh7 cells. A-C) Graphs showing the relative fluorescence of resorufin of pla B compared to a DMSO control for 24, 4 and 1 hour respectively. **D)** A line graph representing the change in resorufin fluorescence over time. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values represent calculated using a one-way ANOVA compared to pri-miR-122.

4.3: The effect of splice inhibition and Microprocessor knockdown on endogenous pri-miR-122

As shown in Chapter 4.2, pla B can be effectively used to inhibit splicing in Huh7 cells. To understand the effect the Microprocessor has on endogenous pri-miR-122 it was important to determine that siRNA treatment had the desired effect on Huh7 cells. Huh7 cells were treated with control, Drosha or DGCR8 siRNA for 72 hours and the proteins harvested in RIPA buffer. As expected, both Drosha and DGCR8 siRNA reduced the reciprocal proteins to undetectable levels (**Figure 4.11**). Drosha siRNA treatment also lead to an increase in the levels of DGCR8 protein, while DGCR8 siRNA lead to Drosha also being knocked down; this has been described in the literature^{116,124,125}.

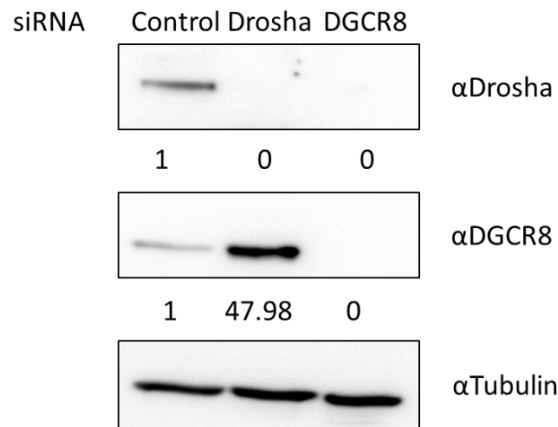


Figure 4.11: DGCR8 knockdown leads to a reduction in DGCR8 and Drosha protein expression. Western blot analysis of the Microprocessor proteins following siRNA treatment. The numbers below the blot are quantifications of the proteins detected.

To understand how splice inhibition and the Microprocessor may interact, Huh7 cells that had been treated with siRNA targeting one of the Microprocessor proteins, DGCR8 or Drosha, were then treated with 1 μ M pla B for 4 hours before total RNA was extracted from the cells. The RNA was then reverse-transcribed, and the resulting cDNA used for qPCR analysis. The same unspliced and spliced pri-miR-122 primers were used as shown in figure 4.3. In addition, qPCR primers were also used to measure Drosha cleavage and transcriptional termination. Primers spanning the 5' Drosha cleavage site of the miR-122 hairpin are used to measure how much hairpin is retained in the pri-miR-122, with a lower qPCR signal indicating that less hairpin is retained and more Microprocessing has occurred. Primers amplifying downstream of the hairpin, referred to as 3', measure the levels of transcriptional termination, with less qPCR signal indicating more transcriptional termination has occurred (**Figure 4.12**).

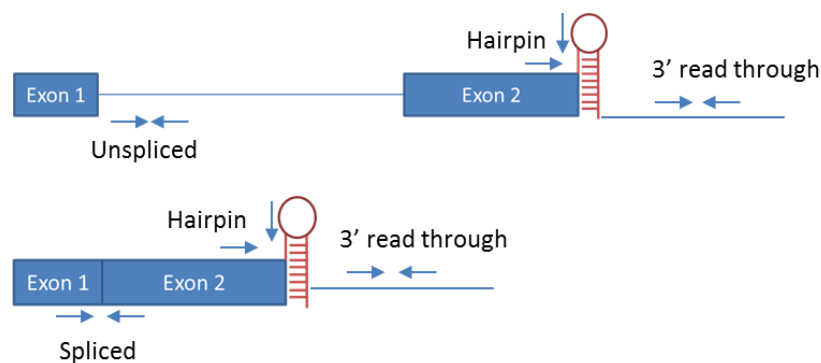


Figure 4.12: A schematic of the qPCR binding sites in pri-miR-122. The blue arrows represent the qPCR binding sites.

When looking at the pri-miR-122 transcript, there are a few ways in which this qPCR data can be analysed: all samples compared to the DMSO treated control siRNA sample; each pla B treated sample compared to the relative DMSO treated sample; DMSO and pla B samples compared to the DMSO or pla B treated control siRNA.

When Huh7 cells are treated with siRNA to knockdown the Microprocessor, the qPCR data shows a possible reduction in the levels of unspliced pri-miR-122 decrease in comparison to the control siRNA treated sample (**Figure 4.13**). When the control siRNA sample is treated with pla B, the levels of unspliced pri-miR-122 decreases as seen in the absence of siRNA treatment, shown in figure 4.5.

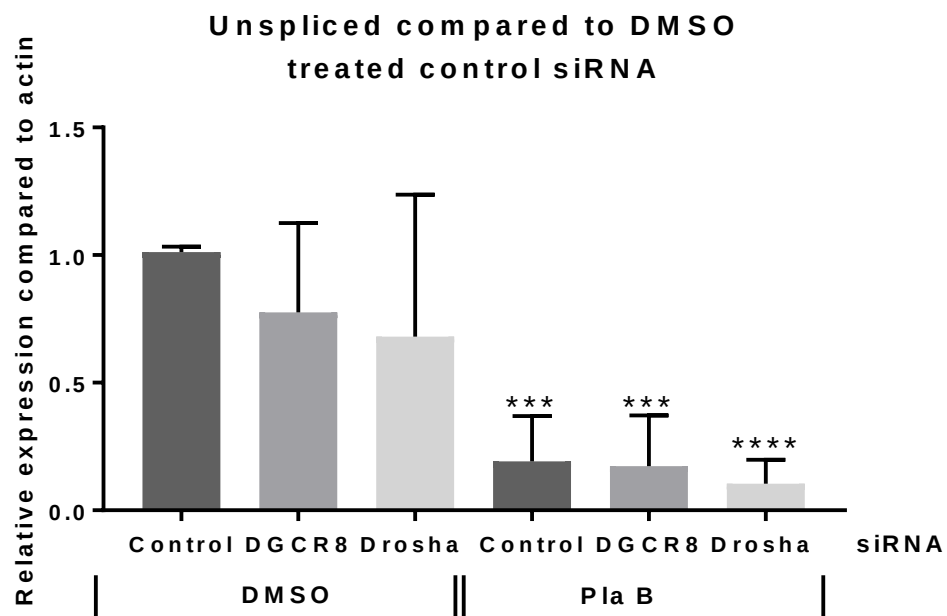


Figure 4.13: Splice inhibition or Microprocessor knockdown affects the levels of endogenous unspliced pri-miR-122. A graph representing the relative expression of unspliced pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated control siRNA sample. Error bars represent standard deviation with a minimum of 6 biological replicates and p-values calculated using a one-way ANOVA compared control DMSO with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

When Huh7 cells are treated with siRNA to knockdown the Microprocessor and pla B to inhibit splicing, the levels of unspliced pri-miR-122 decrease compared to the DMSO treated siRNA control (**Figure 4.13**). When analysing the data in this way, it is clear that the Microprocessor may be having an effect on the levels of spliced pri-miR-122 although it is difficult to

ascertain what effect Microprocessor knockdown is having on unspliced pri-miR-122 levels in conjunction with pla B.

Another way of analysing this data is to compare each pla B treated sample to the DMSO treated sample for the same siRNA treated sample. When analysing the data in this way, splice inhibition is shown to reduce the levels of unspliced pri-miR-122 for all samples (**Figure 4.14**).

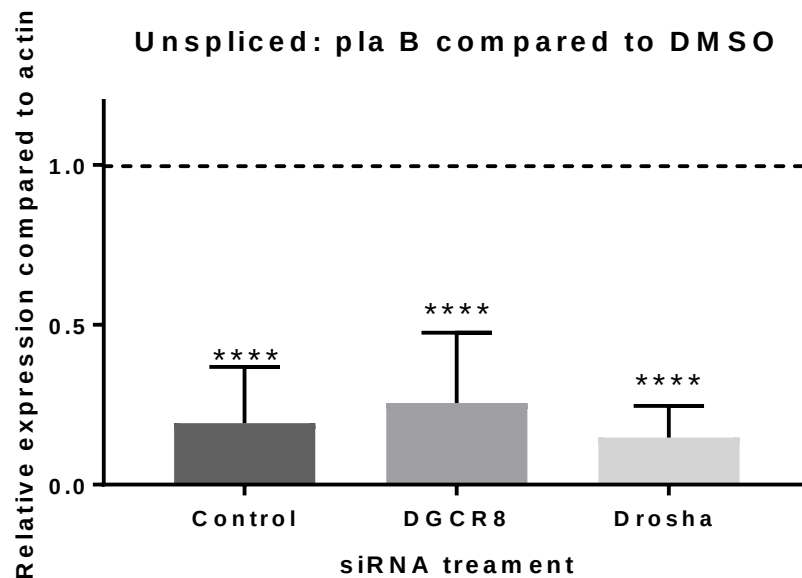


Figure 4.14: Splice inhibition in conjunction with Microprocessor knockdown may affect the levels of endogenous unspliced pri-miR-122. A graph representing the relative expression of unspliced pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to an actin control. Each pla B treated sample has been compared to the DMSO treated sample of the same siRNA. Error bars represent standard deviation with a minimum of 6 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

When considering that both unspliced and spliced levels of pri-miR-122 decrease when Huh7 cells are treated with pla B to inhibit splicing, it is perhaps not surprising that qPCR data for spliced pri-miR-122 shows very similar results to the unspliced (**Figure 4.15**). However, while there is still a significant decrease in the levels of spliced pri-miR-122 when Huh7 cells are treated with siRNA to knockdown the Microprocessor, there is less of a decrease when compared to unspliced transcript in figure 4.15. This decrease in spliced pri-miR-122 following Microprocessor knockdown may indicate that the loss of the Microprocessor is affecting transcription, although more data is necessary to draw a strong conclusion.

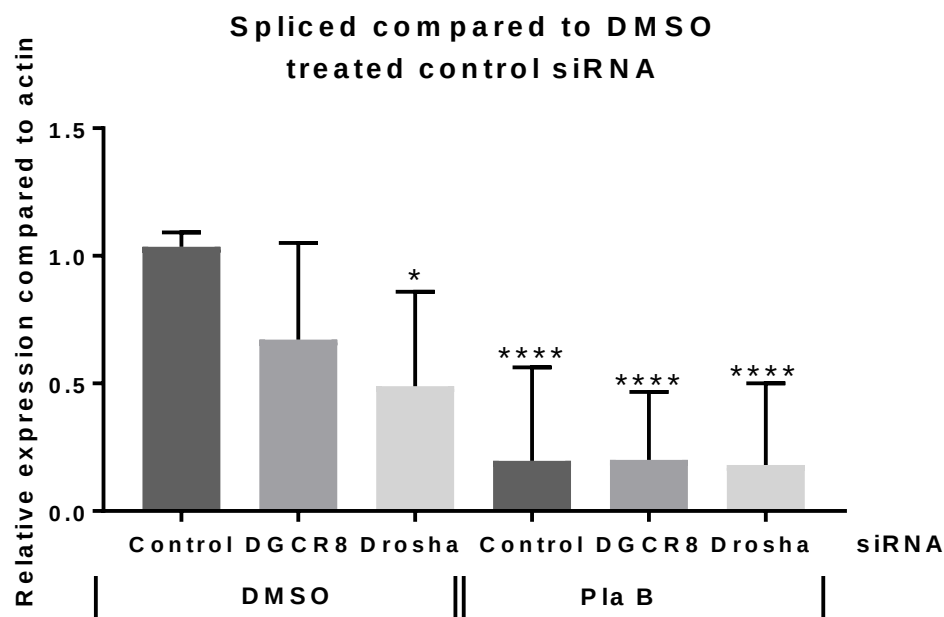


Figure 4.15: Splice inhibition or Microprocessor knockdown affects the levels of endogenous spliced pri-miR-122. A graph representing the relative expression of spliced pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated control siRNA sample. Error bars represent standard deviation with a minimum of 7 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

When the pla B samples are shown relative to their DMSO control, there is a significant decrease in the levels of spliced pri-miR-122 in all conditions as expected (**Figure 4.15**). When the DGCR8 siRNA treated sample is compared to the control siRNA treated sample, the qPCR data shows that there was a significant increase in the level of spliced pri-miR-122. This is also shown in the previous data looking at unspliced pri-miR-122 (**Figure 4.14**). This data could suggest the Microprocessor cleavage is required for splicing and efficient transcription. Seeing that DGCR8 siRNA treatment with pla B has less of an effect on both spliced and unspliced pri-miR-122 indicates that this could be a true effect although it may not be biologically relevant as the effect is quite small.

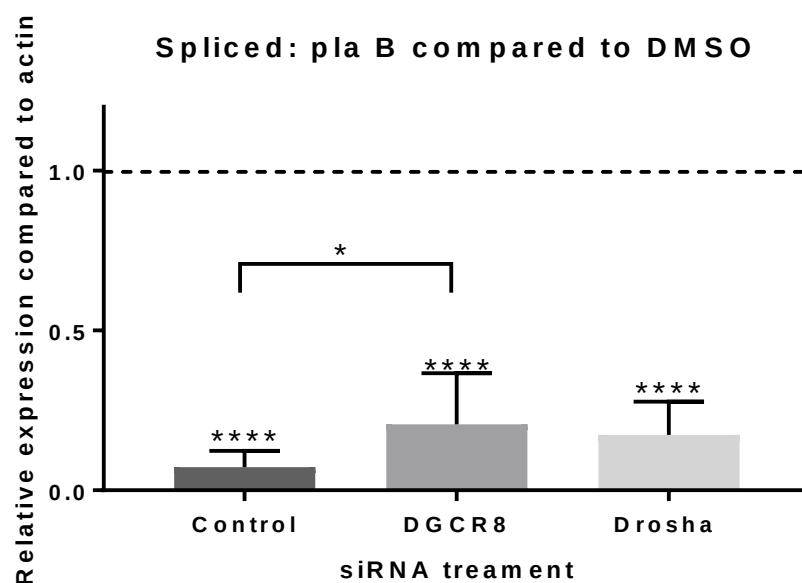


Figure 4.16: Splice inhibition in conjunction with Microprocessor knockdown has little effect on the levels of endogenous spliced pri-miR-122. A graph representing the relative expression of spliced pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to an actin control. Each pla B treated sample has been compared to the DMSO treated sample of the same siRNA. Error bars represent standard deviation with a minimum of 7 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

To understand the effects of splice inhibition and Microprocessor knockdown on miRNA biogenesis, qPCR primers amplifying a region spanning the 5' Drosha cleavage site were used. The more signal detected during qPCR analysis, the more the hairpin is being retained in the pri-miR-122 transcript, indicating less miRNA biogenesis are occurring. When either of the Microprocessor proteins are knocked down, the levels of hairpin signal are significantly increased, as expected (**Figure 4.17**). This indicates that Microprocessor knockdown is affecting miR-122 biogenesis.

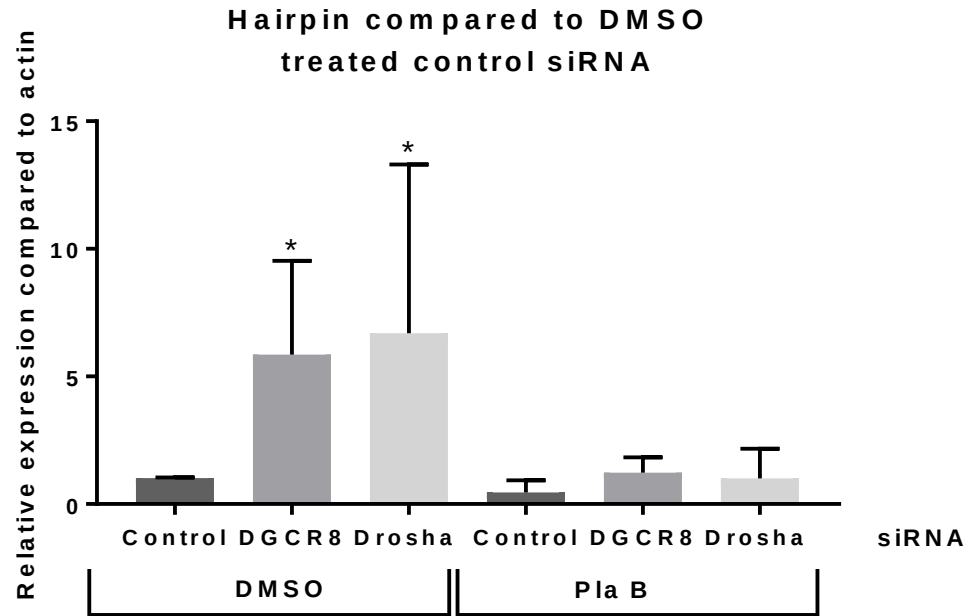


Figure 4.17: Microprocessor knockdown leads to a decrease in miR-122 processing. A graph representing the relative expression of miR-122 hairpin retained in the pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated control siRNA sample. Error bars represent standard deviation with a minimum of 6 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$.

When the cells are treated with pla B to inhibit splicing, the levels of miR-122 hairpin retained in the pri-miR-122 transcript is significantly decreased for the control and Microprocessor knockdown samples (**Figure 4.18**). This fits with the previous observation that unspliced pri-miR-122 is reduced with pla B treatment (**Figure 4.14**).

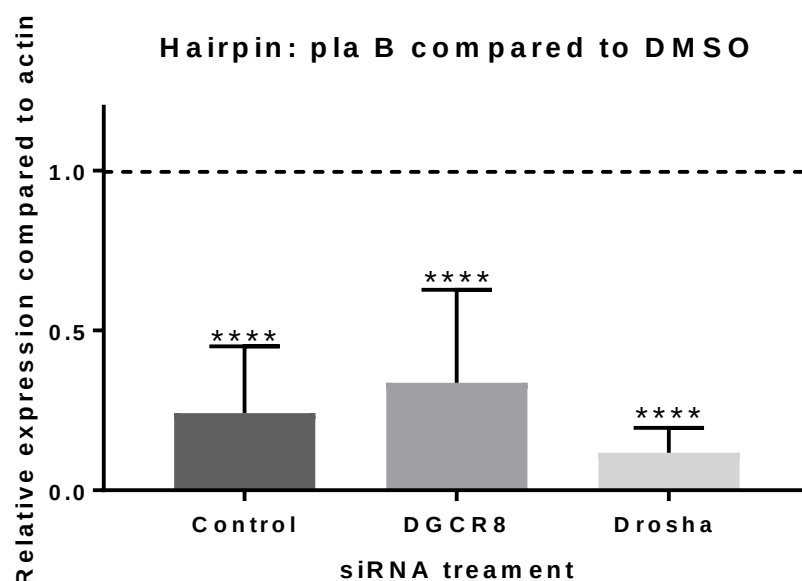


Figure 4.18: Splice inhibition with Drosha knockdown affects the levels of endogenous pri-miR-122 hairpin compared to pla B treated siRNA control. A graph representing the relative expression of miR-122 hairpin in the pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to an actin control. Each pla B treated sample has been compared to the DMSO treated sample of the same siRNA. Error bars represent standard deviation with a minimum of 5 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

Pri-miR-122 transcription is terminated via the Microprocessor, rather than the CPA (Cleavage and Polyadenylation) mechanism and Microprocessor knockdown has been shown to increase the level of transcriptional read-through indicating inefficient transcriptional termination²⁰⁶. The data shown in figure 4.19, shows that transcriptional read-through is significantly increased when DGCR8 is knocked down, however little effect is detected when the other Microprocessor protein, Drosha, is knocked down. This lack of transcriptional read-through seen during Drosha knockdown may be explained by the variation in the samples which could be rectified by more repeats. When control or Microprocessor siRNA treated Huh7 cells were treated with pla B to inhibit splicing the levels of transcription read-through significantly decreases.

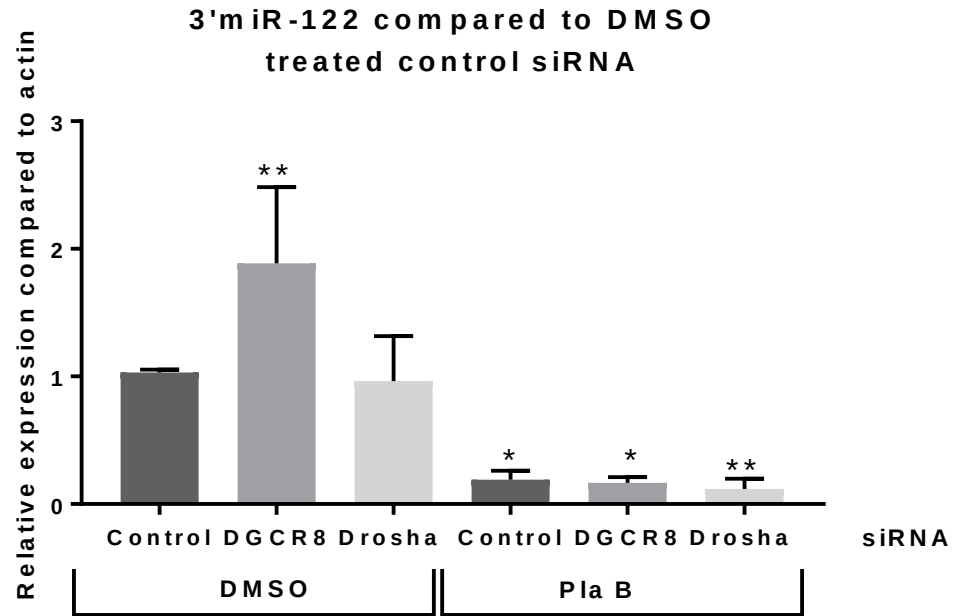


Figure 4.19: Splice inhibition leads to an increase in pri-miR-122 transcriptional read through. A graph representing the relative expression of miR-122 read through in the pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated control siRNA sample. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

With splice inhibition being shown to decrease the levels of spliced and unspliced pri-miR-122 by reducing the levels of transcription, a decrease in transcriptional read through as shown in figure **4.19** and **4.20** would be expected. DGCR8 and Drosha siRNA has no additional effect when combined with pla B treatment on the level of transcriptional read through compared to control siRNA (**Figure 4.20**).

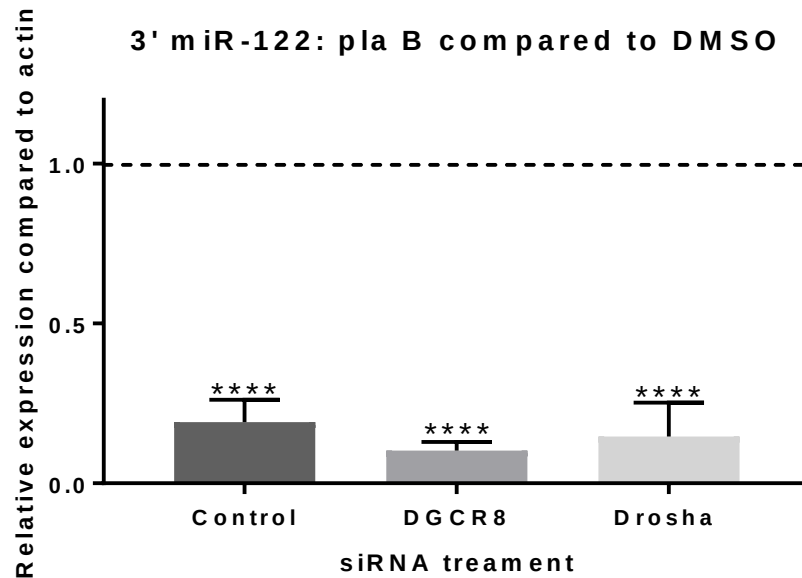


Figure 4.20: Splice inhibition has no additional effect on transcriptional termination efficiency of pri-miR-122 with Microprocessor knockdown. A graph representing the relative expression of pri-miR-122 read-through in Huh7 cells analysed by qPCR and compared to an actin control. Each pla B treated sample has been compared to the DMSO treated sample of the same siRNA. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

qPCR analysis was used to measure the levels of mature miR-122 from Huh7 cells following Microprocessor knockdown and pla B treatment (**Figure 4.21**). While there is a lot of variability within the repeats, the data indicates that the levels of mature miR-122 are relatively stable compared to the control. As miR-122 is known to be a very stable miRNA this is perhaps not surprising²³⁸, although after 72 hours a decrease may be expected, especially when considering there is a marked decrease in pri-miR-122 expression following pla B treatment. An effect on the levels of mature miR-122 expression may be masked by the high variability within the data.

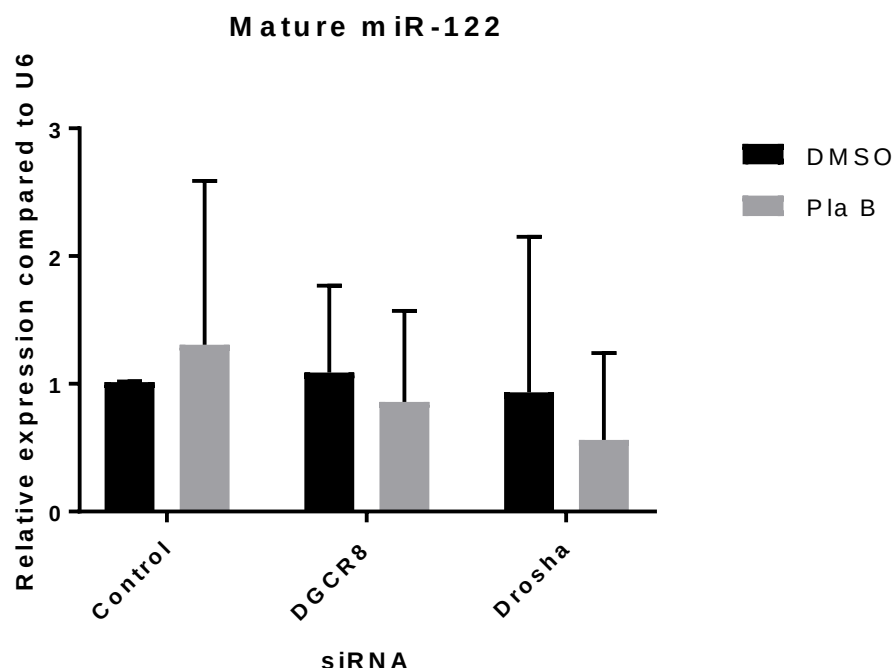


Figure 4.21: Splice inhibition using pla B and Microprocessor knockdown has little effect on the levels of endogenous mature miR-122. A graph representing the relative expression of mature miR-122 in Huh7 cells analysed by qPCR and compared to a U6 control. All samples have been compared to the DMSO treated with the same siRNA. Error bars represent standard deviation with a minimum of 2 biological replicates and p-values calculated using a one-way ANOVA compared to pri-miR-122.

4.4: SF3b1 siRNA treatment has a similar effect on endogenous pri-miR-122 as pla B

To ensure that the effects of pla B on pri-miR-122 are due to the effects of splice inhibition, SF3b1 siRNA was used.

Huh7 cells treated with SF3b1 siRNA for 72 hours were harvested, and the levels of total spliced and unspliced pri-miR-122 were analysed by qPCR. The levels of Drosha cleavage and transcriptional termination were also analysed by qPCR using the primers indicated in figure 4.12. In addition, the levels of unspliced pri-miR-17~92a, pri-miR-23a~27a~24-2 and pri-miR-29a~29b-1, which are also located within lncRNAs, were analysed by qPCR.

As anticipated, the levels of spliced and unspliced pri-miR-122 decreases when SF3b1 is knockdown in Huh7 cells. Interestingly, the relative expression of the miR-122 hairpin and 3' read through also decrease compared to the control siRNA samples (**Figure 4.22**).

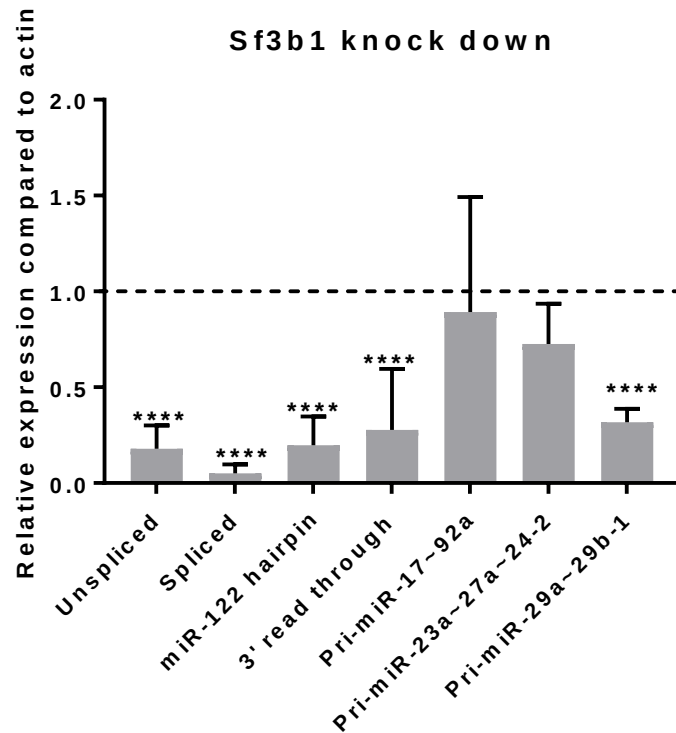


Figure 4.22: Splice inhibition using SF3b1 siRNA reduces the levels of endogenous spliced and unspliced pri-miR-122. A graph representing the relative expression of spliced pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to an actin control. The dotted line represents the siRNA control which has been set to 1. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA compared to pri-miR-122 with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

The levels of pri-miR-23a~27a~24-2 and pri-miR-29a~29b-1 also show a slight decrease in relative expression when Huh7 cells were treated with SF3b1 siRNA. It is also likely that splice inhibition leads to a reduction in pri-miR-17~92a, but due to variations within the samples this effect is hidden. The relative expressions for the three pri-miR-17~92a replicates are: 0.55, 0.53, and 1.57. Using a Grubbs outlier test ($\alpha = 0.1$) the 1.57 value is detected to be an outlier. With this value removed, pri-miR-17~92a is significantly reduced when SF3b1 is knocked down, however this reduces the number of repeats to 2 (**Figure 4.23**).

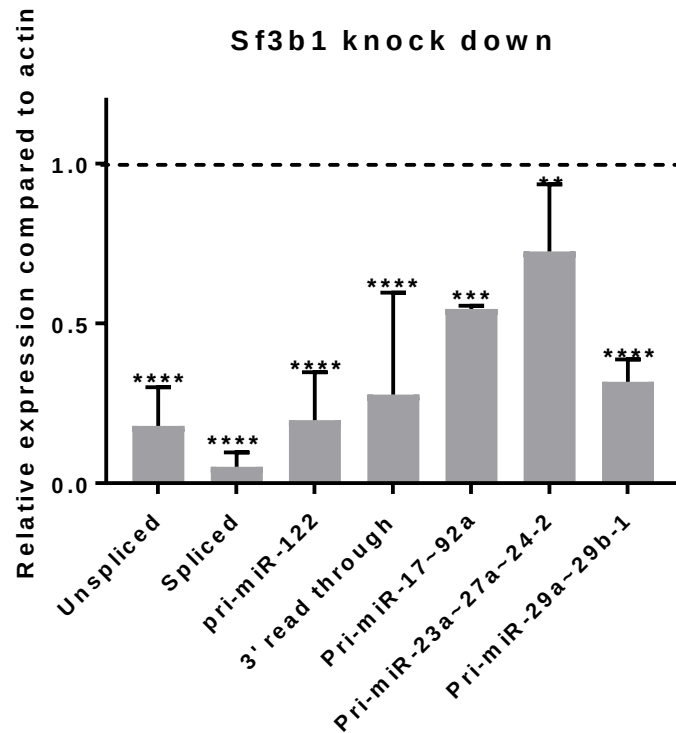


Figure 4.23: Outlier removal shows that splice inhibition using SF3b1 siRNA reduces the levels of different unspliced lnc-pri-miRNAs. A graph representing the relative expression of spliced pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to an actin control. The dotted line represents the siRNA control which has been set to 1. Error bars represent standard deviation with a minimum of 2 biological replicates and p-values calculated using a one-way ANOVA compared to pri-miR-122 with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

As the siRNA treatment used is for 72 hours, should splice inhibition by SF3b1 siRNA effect mature miR-122 expression, this would be detectable by qPCR. The qPCR data suggests that there may be a small increase in the levels of mature miR-122 following SF3b1 siRNA treatment, as opposed to the decrease that would be expected given the decrease in pri-miR-122 levels. However this is not statistically significant and more repeats would be needed to confirm (**Figure 4.24**).

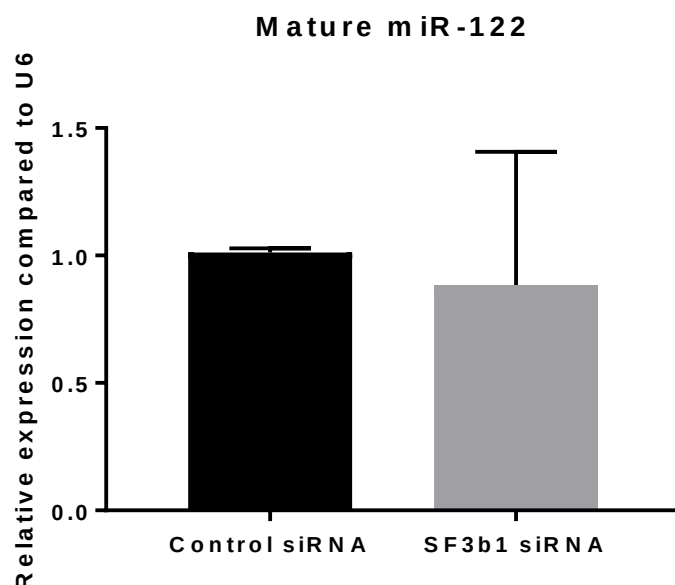


Figure 4.24: The effect of splice inhibition on mature miR-122 following SF3b1 siRNA treatment. A graph representing the relative expression of spliced pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to a U6 control. Error bars represent standard deviation with a minimum of 2 biological replicates and p-values calculated using T-test with * $p < 0.05$.

4.5: The effect of U1 and U2 morpholinos on endogenous pri-miR-122

Morpholinos (AMOs) are synthetic nucleic acids. Unlike DNA and RNA which have a ribose or deoxyribose ring, AMOs have a morpholine ring from which they get their name and each base is linked by an uncharged phosphorodiamidate group instead of a phosphate group. AMOs bind to target sequences and blocks other molecules from interacting with its target and are a well-established method of inhibiting splicing.

Using a β -globin reporter under a HIV-LTR similar to that used in chapters **3**, it has been shown that U1 snRNP interacts with RNA polymerase II independent of splicing, whereas U2 snRNPs are only recruited when a 'functional' intron is present²³⁹. As shown in this chapter, inhibition of splicing by Pla B leads to an increase in the levels of unspliced pri-miR-122 but has the opposite effect on the endogenous pri-miR-122. While treating Huh7 cells with an AMO targeting either U1 will globally reduce splicing and lead to transcriptional instability, when compared to the effects of U2 AMO, this may indicate why endogenous and overexpressed pri-miR-122 is affected differently by splice inhibition.

Huh7 cells were transfected with control, U1 or U2 morpholinos by electroporation and harvested after 24 hours. The RNA was extracted and the levels of spliced and unspliced pri-miR-122 was analysed by qPCR. With only one repeat of this experiment successful, it is difficult to draw meaningful conclusions from this data. The data does suggest that U1 AMO treatment does lead to a stronger decrease in spliced pri-miR-122 and also has a strong effect on the levels of spliced pri-miR-122 (**Figure 4.25**). Similarly, U2 AMO treatment also leads to a decrease in the levels of spliced and unspliced pri-miR-122 as would be expected based on the Pla B and SF3b1 siRNA data shown previously. As it is known that U1 snRNP knockdown leads to transcriptional instability, this data could suggest that the loss of U2 snRNP may also lead to unstable transcription.

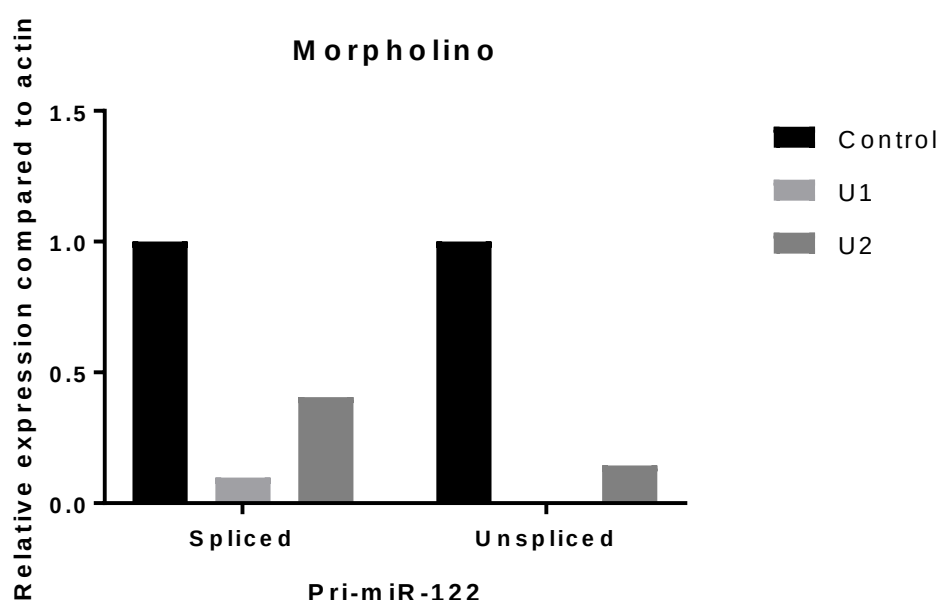


Figure 4.25: U1 and U2 AMOs lead to a reduction in spliced and unspliced pri-miR-122. A graph representing the relative expression of spliced and unspliced pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to an actin control. No error bars are shown as this is a single repeat.

4.6: Discussion

The work in this chapter has shown that treating Huh7 cells, and HeLa cells transfected with pri-miR-122, with pla B is an effective method of inhibiting splicing. At 4 hours of treatment with 1 μ M pla B the levels of spliced pri-miR-122 reduces for both the endogenous and overexpressed transcript. Interestingly, splice inhibition has different effects on the levels of endogenous unspliced pri-miR-122 expressed in Huh7 cells compared to

the levels found in HeLa cells overexpressing pri-miR-122. Treating pri-miR-122-transfected HeLa cells with pla B reduced the levels of splice pri-miR-122 while the levels of the unspliced transcript increases as expected. While pla B reduced the levels of endogenous spliced pri-miR-122 in Huh7 cells, the unspliced transcript similarly decreases. This data may indicate that splice inhibition with pla B may be negatively affecting endogenous pri-miR-122 transcription.

To ensure that the effect on endogenous unspliced pri-miR-122 was not unique to pla B, two other methods of splice inhibition were used. When using siRNA targeting the U2 snRNP subunit SF3b1, both unspliced and spliced pri-miR-122 levels decreased. However, it is worth noting that a 72 hour knockdown protocol was used which resulted in a lot of cell death. This is not surprising as a 72 hour knockdown will likely have a significant effect on gene expression. Longer treatment times will likely to cause secondary effects on gene expression which may affect the data. While the effect of SF3b1 knockdown is similar to the effect of pla B treatment, pla B is preferable due to the shorter treatment time.

The effect of SF3b1 knockdown was also used to look at the unspliced transcripts three other miRNAs located within lncRNAs: pri-miR-17~92a, pri-miR-23a~27a~24-2 and pri-miR-29a~29b-1. This data suggests that the effect of splice inhibition on unspliced pri-miR-122 is not unique, and may extend to other lncRNAs containing miRNA.

While the effects on the spliced pri-miR-122 indicates that splicing has been knocked down by SF3b1 siRNA, a western blot has not be used to confirm the efficiency of the knockdown. An attempt to analyse the levels of SF3b1 mRNA by qPCR was unsuccessful. Therefore, while the data obtained from SF3b1 knockdown is similar to the data obtained from pla B treatment, it cannot be confirmed that the effects seen with SF3b1 are due to splicing knockdown.

AMOs targeting the U1 and U2 snRNPs were also used. While the literature suggests that an effect on splicing can be seen within 8 hours, no effect on splicing was seen in this time²⁴⁰. At 24 hours, both U1 and U2 AMOs lead to a reduction in spliced and unspliced endogenous pri-miR-122, with U1 showing a stronger reduction. This is not surprising as the U1 snRNP

associates with RNA polymerase II independently to splicing²⁴¹. As U2 AMO treatment has a similar effect to U1, this may indicate that splice inhibition may negatively impact transcription of endogenous miR-122. Due to time constraints, only one repeat of the morpholino experiment was obtained.

Splice inhibition was then compared with Microprocessor knockdown to understand what effect both these processes have on endogenous pri-miR-122. As expected, when Huh7 cells were treated with pla B, the levels of spliced and unspliced pri-miR-122 decreases, with a stronger decrease seen in the levels of spliced transcripts. Also expected, a decrease occurred in the levels of unspliced pri-miR-122 with Microprocessor knockdown which has been shown in the literature²⁰⁶. It has also been shown that the Drosha cleavage decrease and transcriptional read through increases with Microprocessor knockdown which is also shown in this data, although, the increase in 3' read through was not as high as expected²⁰⁶.

When looking at qPCR data for spliced and unspliced pri-miR-122, Microprocessor knockdown does not seem to have any additional effect on the transcripts when combined with pla B treatment. Similarly, the levels of miR-122 hairpin retained in the transcripts, and the levels transcriptional read through are reduced similar to the levels of unspliced pri-miR-122, and no additional effects are seen with Microprocessor knockdown.

Chapter 5: The effect of splice inhibition on chromatin-associated endogenous pri-miR-122

5.1: Introduction

In the previous chapter, it was shown that splice inhibition leads to a decrease in the levels of spliced and unspliced endogenous pri-miR-122 in Huh7 cells. This decrease in unspliced pri-miR-122 could be due to a decrease in transcription. To understand what effect splice inhibition by pla B may be having on the transcription of endogenous pri-miR-122, the next step would be to look at the transcripts as they are being transcribed by RNA polymerase II.

Chromatin-associated RNA and nucleoplasmic RNA were extracted from Huh7 cells, which endogenously express miR-122, using a protocol derived from the literature^{241,246}. Newly transcribed RNA should still be associated with the chromatin and the RNA-polymerase machinery, including the spliceosome and Microprocessor, and should be in the process of being processed by these complexes, while nucleoplasmic RNA has been released from the chromatin. This approach will therefore allow future analysis of co-transcriptional splicing and pri-miRNA processing.

5.2: Chromatin-associated RNA

To ensure the chromatin-associated RNA protocol efficiently separates chromatin from the nucleoplasm, and also the nucleoplasm from the cytoplasm, protein samples were analysed by western blot. The blot indicates that the protocol is cleanly separating the three fractions. As a major component of the cytoskeleton, an antibody binding to tubulin was used to detect the cytoplasm, and figure **5.1** shows that tubulin is only detected in the cytoplasm, as expected. Lamin A/C are important in the structure of the nucleus, in particular the nuclear envelope, and have been shown to interact with chromatin²⁴³. Using an antibody to detect lamin A/C, the western blot shows lamin A/C to be present in the chromatin, with a small amount also detected in the nucleoplasm. An antibody binding to the histone protein H2B only detected H2B in the chromatin fraction, indicating that only the chromatin fraction contains chromatin.

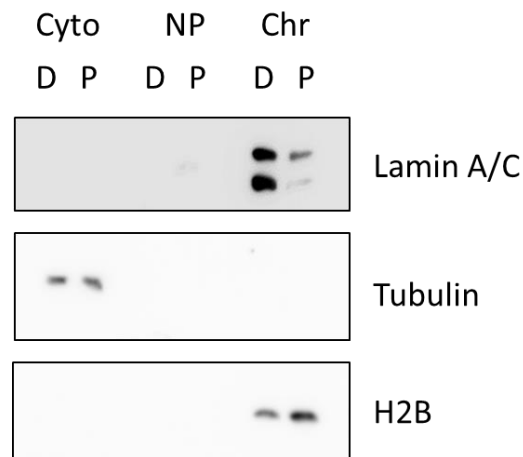


Figure 5.1: Western blot analysis of fractions obtained following the chromatin-associated RNA extraction protocol. Samples were taken from HeLa cells transfected with pri-miR-122 treated with DMSO (D) or pla B (P) at different stages of the protocol to represent cytoplasm, nucleoplasm and chromatin.

While the western blot indicated that the nucleoplasm and the chromatin were efficiently separated, it was important to see if chromatin-associated RNA was fractionated from chromatin-released nucleoplasmic RNA. qPCR was used to analyse the levels of spliced and unspliced endogenous pri-miR-122 transcripts in Huh7 cells as well as with primers across the Drosha cleavage site (hairpin), which indicates the level of miR-122 hairpin retrained in the pri-miRNA, and with primers to detect the level of transcriptional read through (3'). The locations of primer amplification are shown in figure 5.2.

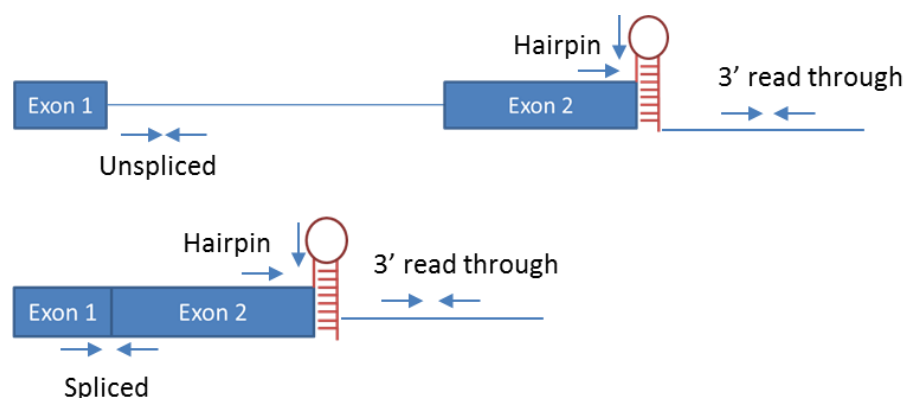


Figure 5.2: A schematic of the qPCR binding sites in pri-miR-122. The blue arrows indicate the regions the qPCR primers bind.

An increase in unspliced pri-miR-122 transcript was observed in the chromatin-associated RNA when compared to nucleoplasmic RNA (**Figure 5.3**). This indicates that RNA extracted from the chromatin has yet to be processed by the spliceosome and is still associated with the RNA

polymerase machinery. Similarly, the levels of miR-122 hairpin and transcriptional read through are also shown to have increased compared to nucleoplasmic RNA. Again, this indicates that the pri-miR-122 is relatively unprocessed compared to nucleoplasmic pri-miR-122 and shows that the protocol is separating nucleoplasmic RNA from chromatin-associated RNA.

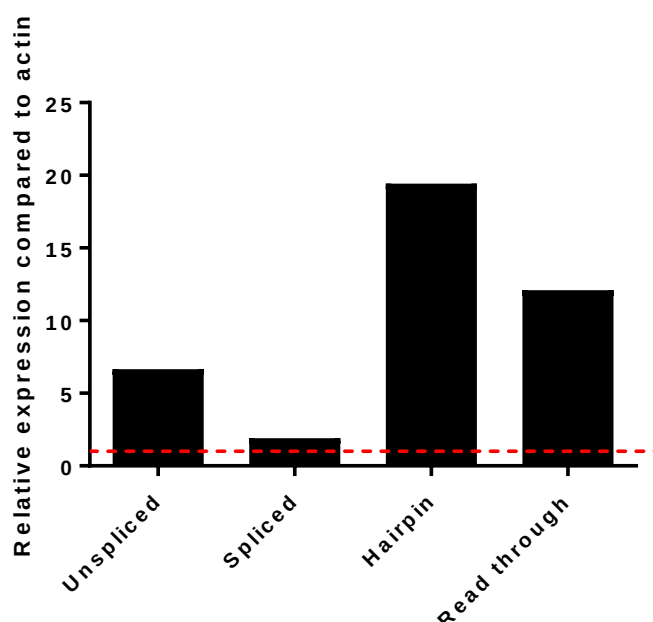


Figure 5.3: Chromatin-associated RNA extraction protocol separates nucleoplasmic RNA from chromatin-associated RNA. A graph representing the relative expression of spliced and unspliced endogenous pri-miR-122, pri-miR-122 read-through and hair-pin levels in Huh7 cells analysed by qPCR and compared to an actin control. All samples have been compared to the nucleoplasmic RNA fraction which is indicated by the red dotted line. No error bar are shown as there is only one replicate.

5.3: The effects of splice inhibition on chromatin-associated pri-miR-122

With the chromatin-associated RNA extraction optimised, the protocol could then be used to analyse the effect of splice inhibition during transcription using pla B. Huh7 cells were treated with 1 μ M pla B for 4 hours, or with DMSO as a control prior to chromatin extraction and qPCR analysis. The primers used in qPCR are outlined in figure 5.2.

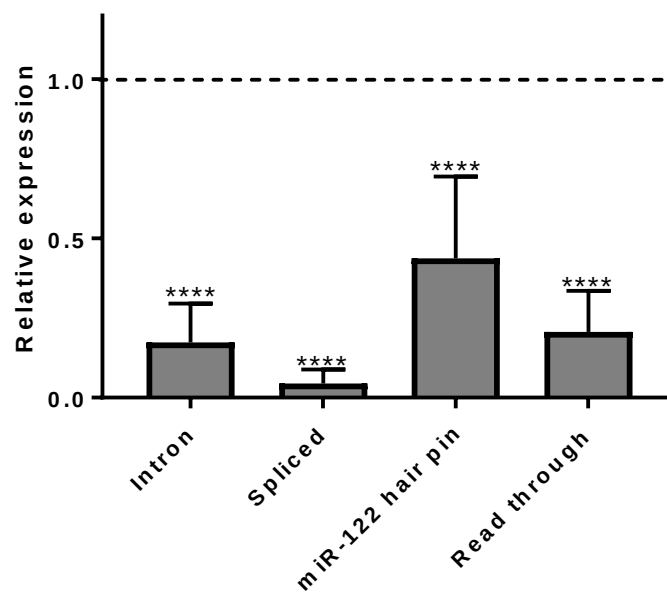


Figure 5.4: Splice inhibition reduces the levels of Chromatin-associated pri-miR-122. A graph representing the relative expression of spliced and unspliced endogenous pri-miR-122, pri-miR-122 read-through and hairpin levels in Huh7 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO-treated chromatin fraction as indicated by the blacked dotted line. Error bars represent standard deviation with a minimum of 6 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

The data indicates that when treated with pla B to inhibit splicing, the levels of spliced and unspliced endogenous pri-miR-122 in the chromatin fraction significantly decrease (**Figure 5.4**). This effect is also seen looking at total RNA from pla B treated huh7 cells discussed in the previous chapter, however we see a greater decrease in figure 5.4 which further supports the conclusion that this effect is happening at the level of transcription. When treated with pla B, the levels of transcriptional read through (3') and the levels of hairpin in the pri-miR-122 (Hairpin) transcript decreases. However, the reductions in the hairpin transcript are less strong than the reduction in unspliced pri-miR-122. If splice inhibition had no effect on Microprocessing, it could be expected that the effect of pla B on hairpin and read through qPCR products would be similar to the effect on unspliced pri-miR-122. Therefore, this data suggests that splice inhibition may also be inhibiting Microprocessing as the levels of hairpin and 3' read through do not decrease as much as the levels of unspliced pri-miR-122.

5.4: The effect of confluency on splice inhibition for chromatin-associated RNA

Confluency is known to affect miRNA biogenesis¹⁸⁰. Some outlier data was removed from the hairpin data in figure 5.4 following statistical testing. While the cells were plated at the same level of confluency each time, it is possible that small changes in confluency could lead to variability. To understand whether this variability was due to confluency, Huh7 cells were plated to 70 % or 100 % confluency, treated with pla B 16 hours after plating, and the chromatin-associated RNA extracted 4 hours after the start of pla B treatment. The RNA was analysed by qPCR using the same primers as show in figure 5.2.

For both normal confluency and over-confluent cells the levels of unspliced pri-miR-122 decreased following pla B treatment to inhibit splicing (**Figure 5.5**). There is no statistical difference between the two pla B treated samples. This could be due to the amount of variation in the samples, with the three replicates for the normal confluency having the following relative expressions: 0.087, 0.049, 0.88. The latter figure shows little effect of pla B and this may be due to contamination with DNA. Should this figure be removed, the levels of unspliced pri-miR-122 for the 70 % and over confluent Huh7 cells would be similar.

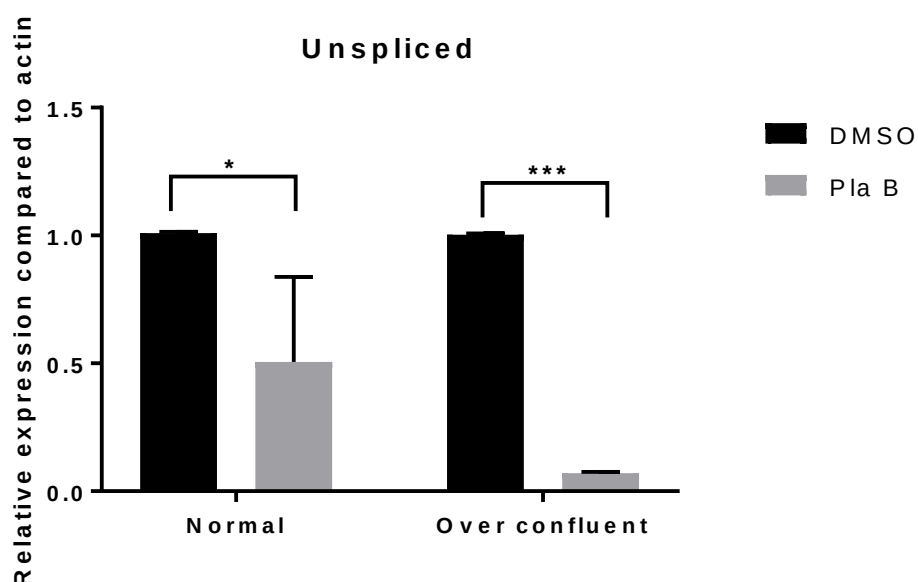


Figure 5.5: Confluency has little effect on chromatin-associated unspliced pri-miR-122 following splice inhibition. A graph representing the relative expression of unspliced pri-miR-122 in Huh7 cells following pla B treatment analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated cells of the same confluency. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values

calculated using a two-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

As expected, the levels of spliced pri-miR-122 decrease significantly following pla B treatment to inhibit splicing for 70 % confluent Huh7 cells (**Figure 5.6**). Over-confluent Huh7 cells treated with pla B decrease to a similar level. This indicates that the confluency does not affect how efficiently splicing is inhibited.

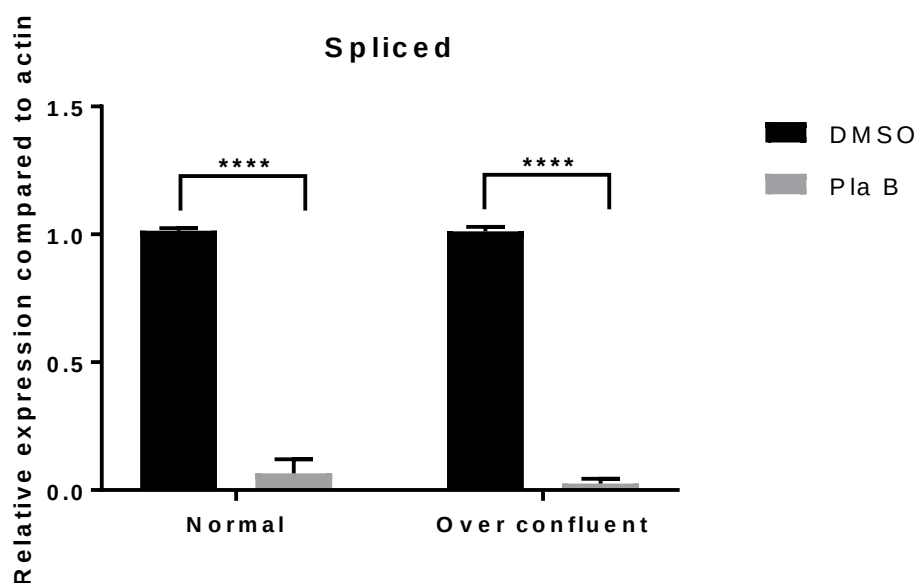


Figure 5.6: Confluency has little effect on chromatin-associated spliced pri-miR-122 following splice inhibition. A graph representing the relative expression of spliced pri-miR-122 in Huh7 cells following pla B treatment analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated cells of the same confluency. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

In agreement with the data shown in figure **5.4**, at normal confluency, the levels of hairpin decrease following pla B treatment (**Figure 5.7**). The hairpin primers span the Drosha cleavage site and are an indication of how efficiently the Microprocessor is processing the miR-122 hairpin. When over-confluent Huh7 cells are treated, the hairpin levels detected by qPCR also decrease compared to the DMSO control. Interestingly, the level of hairpin also decreases significantly in the over-confluent cells compared to normally confluent cells. This data suggests that over-confluent Huh7 cells are more efficiently processing the miR-122 hairpin following splice inhibition with pla B.

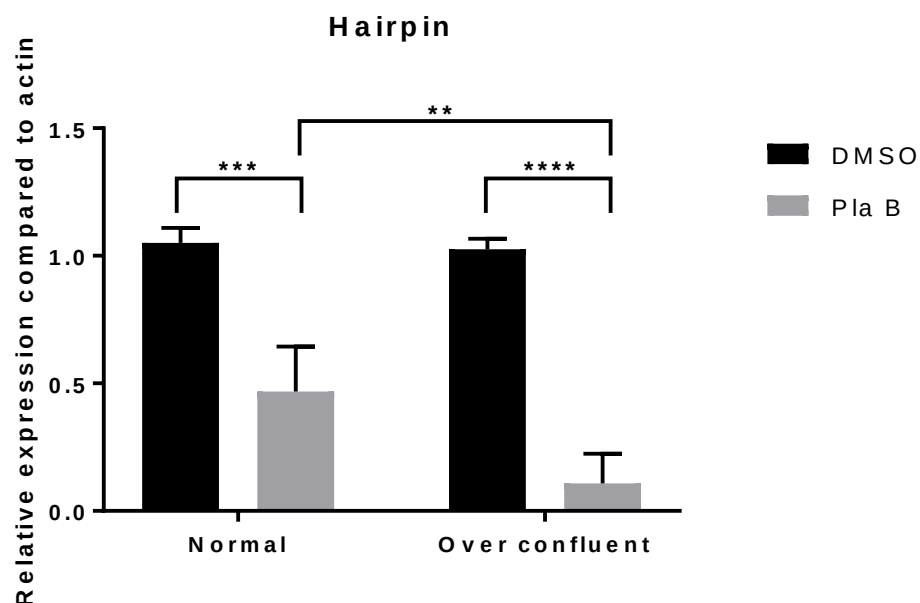


Figure 5.7: Confluency may affect Drosha cleavage efficiency following splice inhibition. A graph representing the relative expression of pri-miR-122 hairpin in Huh7 cells following pla B treatment analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated cells of the same confluency. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

As with spliced and unspliced pri-miR-122 confluency has little effect on the level of 3' read through, and 3' read through similarly decreases for cells at both confluencies (**Figure 6.8**). This data indicates that confluency has no effect of transcriptional termination following pla B treatment to inhibit splicing.

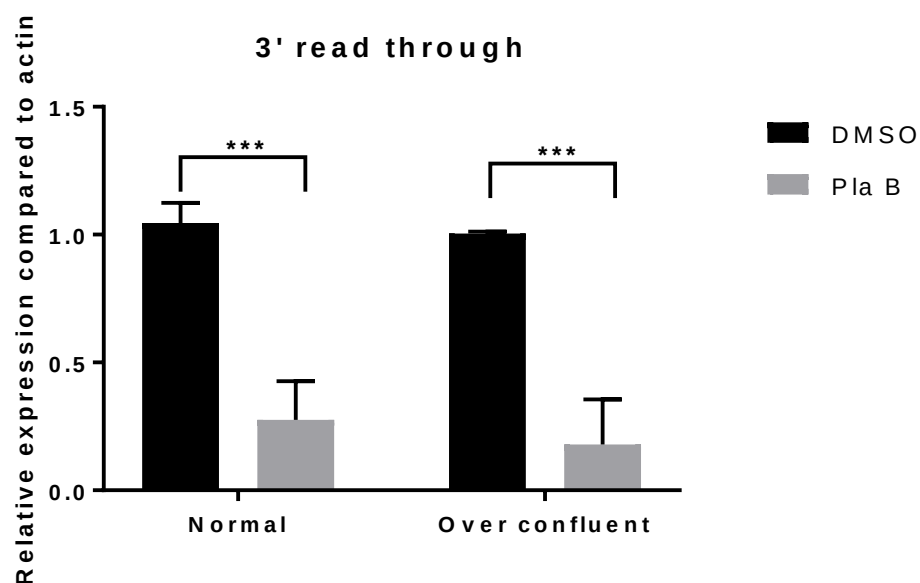


Figure 5.8: Confluency has little effect on chromatin-associated pri-miR-122 3' read through following splice inhibition. A graph representing the relative expression of pri-miR-122 read through in Huh7 cells following pla B treatment analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated cells of the same confluency. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.005$.

5.5: Splice inhibition decreases the transcription of some, but not all, lnc-pri-miRNAs

To understand whether the effect of splice inhibition is unique to chromatin-associated pri-miR-122, qPCR was used to analyse the unspliced host transcripts of three other miRNAs that are expressed from a lncRNA transcript (**Figure 5.9**). The data indicates that pri-miR-29a~29b-1 is not affected by splice inhibition using pla B. However, significant decreases in unspliced transcripts were detected for pri-miR-17~92a and pri-miR-23a~27a~24-2. Interestingly, the levels of unspliced pri-miR-17~92a decrease similarly to unspliced pri-miR-122. This indicates that splice inhibition may affect a subset of miRNAs encoded within lncRNAs.

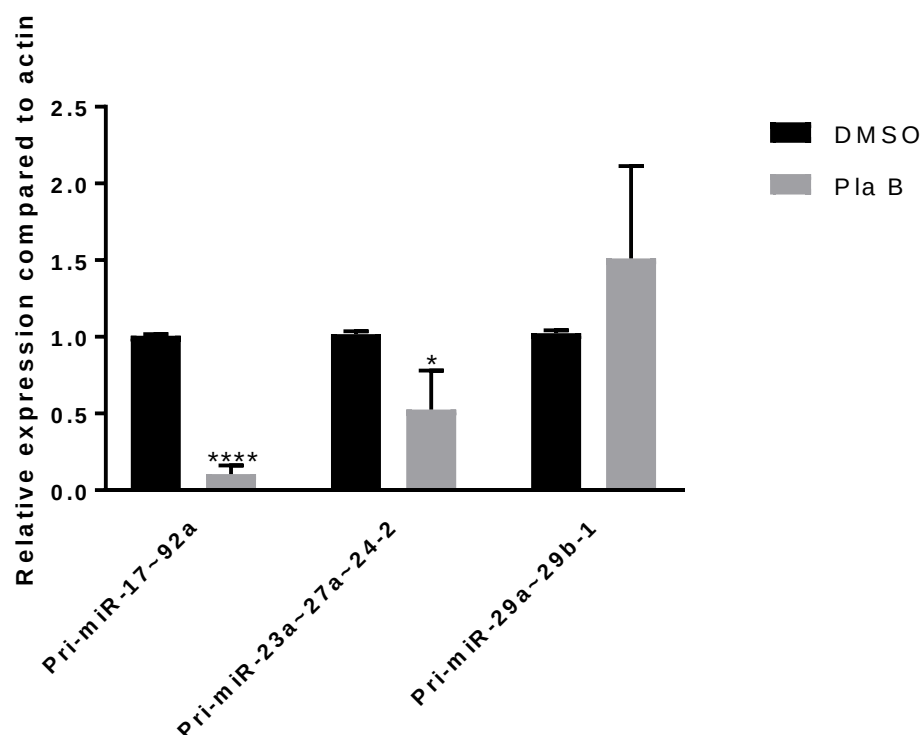


Figure 5.9: Splice inhibition leads to a reduction of some unspliced lncRNA containing miRNA hairpins. A graph representing the relative expression of unspliced pri-miR-17~92a, pri-miR-23a~24~2~27a and pri-miR-29a~29b-1 in Huh7 cells following pla B treatment analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated cells of the same confluency. Error bars represent standard deviation with a minimum of 6 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

To determine whether confluency had an effect on the levels of unspliced pri-miR-17~92a, pri-miR-23a~27a~24~2 and pri-miR-29a~29b-1, 70 % confluent and over-confluent Huh7 cells were treated with pla B and chromatin-associated RNA was extracted and analysed by qPCR. At normal confluency, the levels of unspliced pri-miR-17~92a decrease (**Figure 5.10**). However, this decrease is not as strong as expected and is not statistically significant. When looking at the replicates, one figure of the three shows no effect of pla B treatment. While this could indicate that there was an error during pla B treatment, this effect is not seen for pri-miR-122 for the same RNA sample. Therefore, is it possible that the RNA sample was contaminated with pri-miR-17~92a DNA. The levels of pri-miR-17~92a decrease significantly compared to control for the over

confluent Huh7 cells, and there is no significant difference between the normal and over-confluent pla B treated samples.

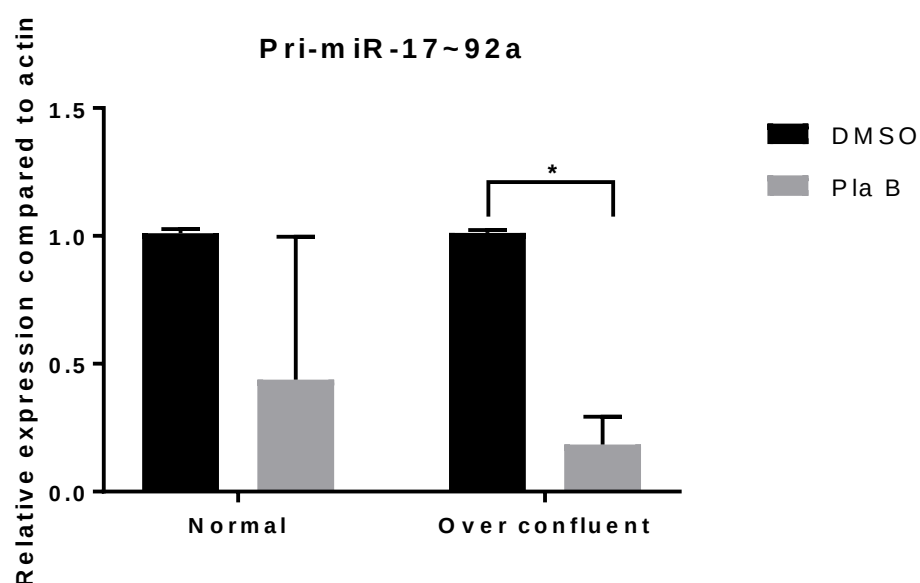


Figure 5.10: Confluency has little effect on chromatin-associated unspliced pri-miR-17~92a following splice inhibition. A graph representing the relative expression of unspliced pri-miR-17~92a in over confluent and 70 % confluent Huh7 cells following pla B treatment analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated cells of the same confluency. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a two-way ANOVA with *p< 0.05.

The same RNA samples as used in the previous figure were also analysed for the levels of unspliced pri-miR-23a~24-2~27a (**Figure 5.11**). The data suggests the pla B treatment has no effect on the levels of unspliced pri-miR-23a~27a~24-2 which is in disagreement with the data shown in figure **5.9**. This data also suggests that confluency has no effect on the levels of unspliced pri-miR-23a~27a~24-2. However, there is a lot a variability between repeats and this could be masking the effect of pla B or the effect of confluency on the levels of unspliced pri-miR-23a~27a~24-2.

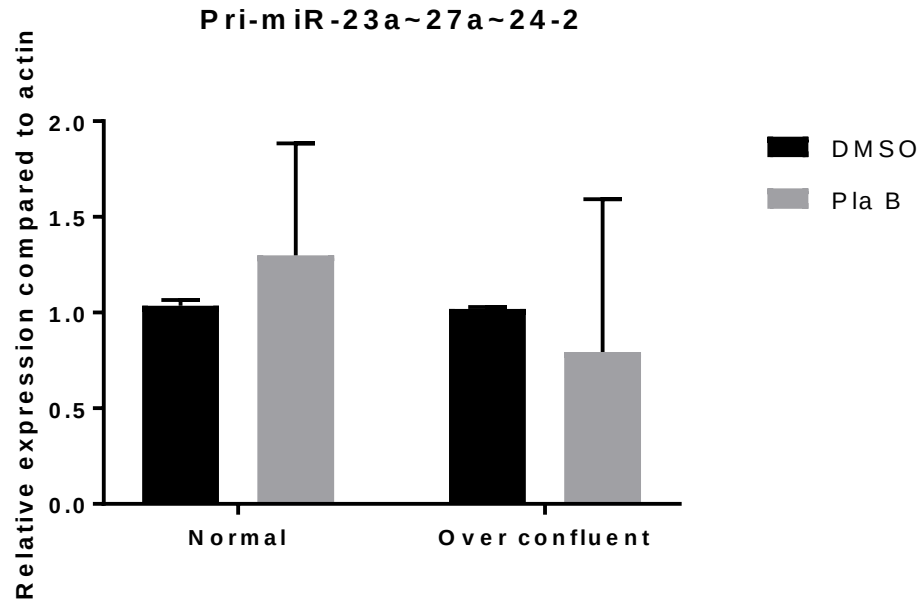


Figure 5.11: Confluency has little effect on chromatin-associated unspliced pri-miR-23a~24-2~27a following splice inhibition. A graph representing the relative expression of unspliced pri-miR-23a~24-2~27a in over confluent and 70 % confluent Huh7 cells following pla B treatment analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated cells of the same confluency. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a two-way ANOVA with *p< 0.05, **p< 0.01, ***p< 0.005, and ****p< 0.0001.

The levels of unspliced pri-miR-29a~29b-1 possibly increase with pla B treatment for Huh7 cells at 70 % confluency, which is also shown in figure **5.8**. This effect on unspliced pri-miR-29a~29b-1 is similar for the over confluent treated cells (**Figure 5.12**). This data indicates that pla B may lead to a small increase in the levels of unspliced pri-miR-29a~29b-1, and this effect is not altered by confluency.

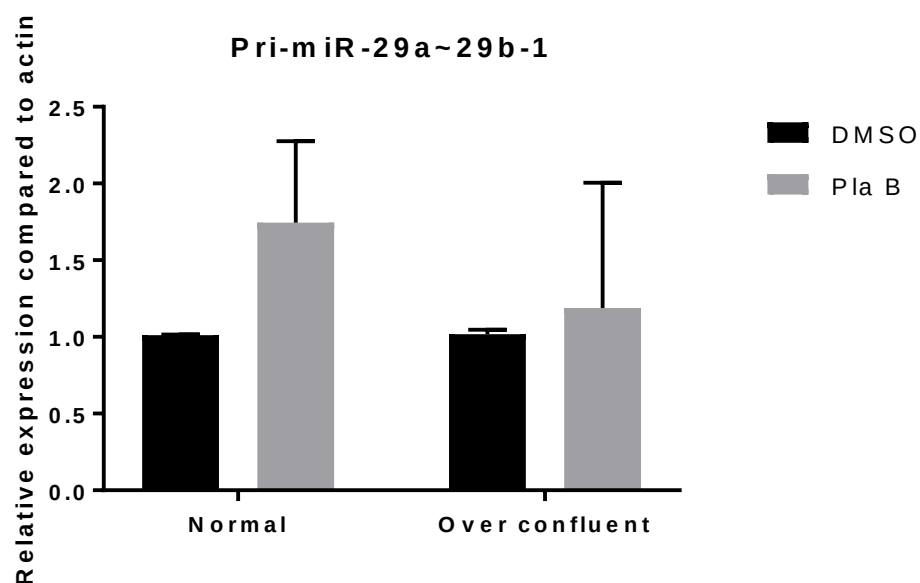


Figure 5.12: Confluency has little effect on chromatin-associated unspliced pri-miR-29a~29b-1 following splice inhibition. A graph representing the relative expression of unspliced pri-miR-29a~29b-1 in over confluent and 70 % confluent Huh7 cells following pla B treatment analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated cells of the same confluency. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a two-way ANOVA.

5.6: The effects of Microprocessor knockdown and splice inhibition on chromatin-associated RNA

In the previous chapter the effect of Microprocessor knockdown in conjunction with pla B treatment to inhibit splicing on pri-miR-122 was analysed. This data shown in this chapter so far also indicates that pla B has a negative effect on transcription as well as inhibiting splicing. To understand whether Microprocessor knockdown influences pla B repression of transcription, Huh7 cells were treated with siRNA to target DGCR8 and then treated with pla B to inhibit splicing before chromatin-associated RNA extraction. Chromatin-associated pri-miR-122 was analysed by qPCR using the primers described in figure 5.1. In addition, the effect of Microprocessor knockdown and splice inhibition on unspliced pri-miR-17~92a, pri-miR-23a~27a~24-2 and pri-miR-29a~29b-1 was also analysed by qPCR.

As expected, and as also shown in figure 5.4, when treated with pla B, the levels of spliced pri-miR-122 decrease compared to DMSO treated control in cells treated with control siRNA, but also for cells treated with DGCR8 siRNA (**Figure 5.13A**). The effect of DGCR8 knockdown on chromatin-

associated spliced pri-miR-122 is variable, with the individual repeats being: 0.55, 0.38, 5.44, 1.10 and 1.16 (**Figure 5.13B**).

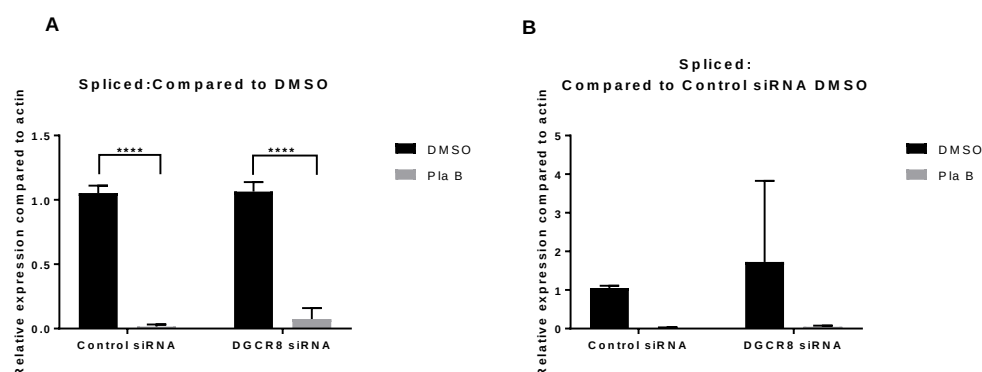


Figure 5.13: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated spliced pri-miR-122. A graph representing the relative expression of spliced pri-miR-122 transcript in Huh7 cells analysed by qPCR and compared to an actin control. **A)** All samples have been compared to the DMSO treated sample for the same siRNA treatment. **B)** All samples have been compared to the DMSO treated control siRNA sample. Error bars represent standard deviation with a minimum of 4 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

The levels of unspliced chromatin-associated pri-miR-122 were analysed by qPCR. As expected, the level of unspliced pri-miR-122 is reduced when treated with pla B, for the control siRNA treated samples, and the DGCR8 siRNA treated samples are similarly decreased (**Figure 5.14A**). There is no significant difference between the pla B samples treated with control or DGCR8 siRNA indicating that the Microprocessor knockdown did not increase or decrease the effects of pla B on transcription. As seen with the spliced pri-miR-122 data, when the Microprocessor is knocked down the levels of DMSO treated chromatin-associated pri-miR-122 varies within repeats (**Figure 5.14B**).

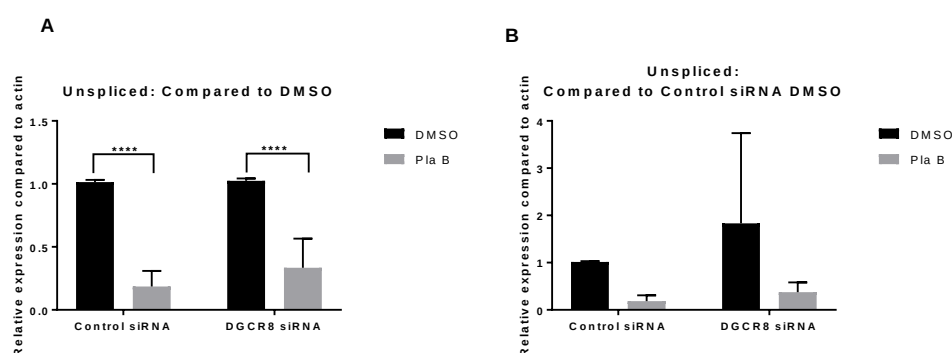


Figure 5.14: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated unspliced pri-miR-122. A graph representing the relative expression of unspliced pri-miR-122

transcript in Huh7 cells analysed by qPCR and compared to an actin control. **A)** All samples have been compared to the DMSO treated sample for the same siRNA treatment. **B)** All samples have been compared to the DMSO treated control siRNA sample. Error bars represent standard deviation with a minimum of 4 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

To understand the effect of Microprocessor knockdown and pla B treatment on Drosha processing, the levels of miR-122 hairpin retained in the pri-miR-122 transcript were analysed by qPCR. When each pla B sample is compared to the relevant DMSO control, the levels of hairpin sequence retained in the pri-miR-122 are significantly reduced (**Figure 5.15A**). A two-way ANOVA shows that the pla B samples are not significantly different. With little to no Microprocessing in these cells, this further indicates that pla B is having a negative effect on transcription and that this effect is not dependent on the Microprocessor.

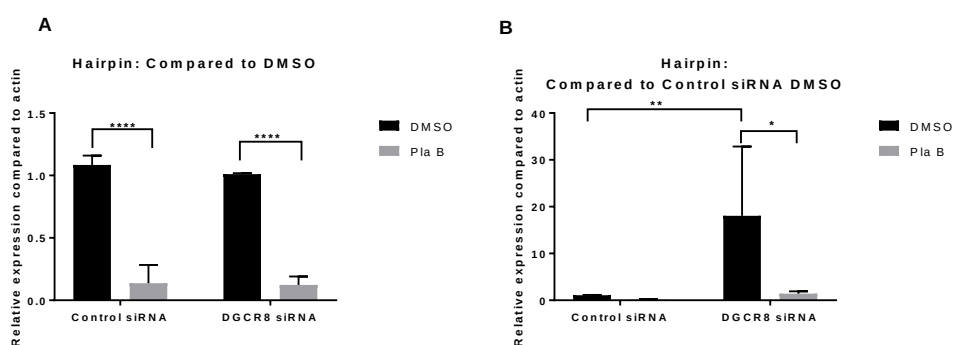


Figure 5.15: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated pri-miR-122 hairpin. A graph representing the relative expression of pri-miR-122 hairpin retained in the transcript in Huh7 cells analysed by qPCR and compared to an actin control. **A)** All samples have been compared to the DMSO treated sample for the same siRNA treatment. **B)** All samples have been compared to the DMSO treated control siRNA sample. Error bars represent standard deviation with a minimum of 4 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$ and **** $p < 0.0001$.

A reduction in Drosha processing following DGCR8 knockdown was detected by qPCR as expected, which is shown by an increase in hairpin signal in DGCR8 siRNA and DMSO treated cells compared to control siRNA DMSO-treated cells in figure 5.14B. The level of hairpin detected does decrease following splice inhibition (**Figure 5.15**) however this result is similar to the decrease seen in the unspliced transcript (**Figure 5.14**). This differs from the previous data where the levels of hairpin retained in the pri-miR-122 transcript increased compared to the level of unspliced pri-miR-122 (**Figure 5.4**).

DGCR8 knockdown is known to lead to an increase in the levels of transcriptional read through, as transcriptional termination is mediated via the Microprocessor²⁰⁶. When pla B treated samples are compared to the relevant DMSO treated samples, the levels of transcriptional read through decrease for both control and DGCR8 siRNA treated cells (**Figure 5.16A**). There is no significant difference between the two pla B treated groups, indicating that the effects of splice inhibition on transcription are not influenced by the Microprocessor.

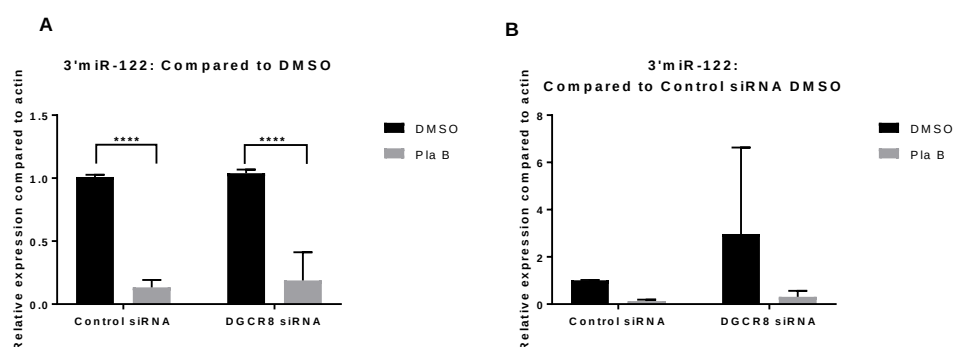


Figure 5.16: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated pri-miR-122 transcriptional read through. A graph representing the relative expression of pri-miR-transcriptional read through in Huh7 cells analysed by qPCR and compared to an actin control. **A)** All samples have been compared to the DMSO treated sample for the same siRNA treatment. **B)** All samples have been compared to the DMSO treated control siRNA sample. Error bars represent standard deviation with a minimum of 5 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

As seen in figure 5.4, pla B treatment reduced the levels of 3' read through detected by qPCR for both the control but also for the DGCR8 treated cells (**Figure 5.16B**). Also, DGCR8 knockdown led to increased transcriptional read through in the DMSO treated cells as expected.

In an earlier figure, the levels of unspliced pri-miR-17~92a were shown to decrease following pla B treatment to inhibit splicing (**Figure 5.10**). This decrease was also seen in the presence of control siRNA in figure 5.17A, and a similar decrease is seen following Microprocessor knockdown and pla B treatment. There is no significant difference between the two pla B treated samples, which indicates that the effects of splice inhibition on unspliced pri-miR-17~92a are not influenced by the Microprocessor.

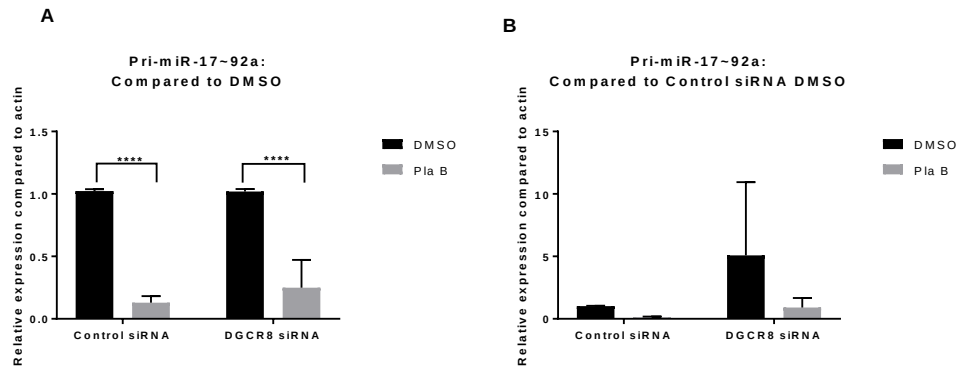


Figure 5.17: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated unspliced pri-miR-17~92a. A graph representing the relative expression of pri-miR-17~92a in Huh7 cells analysed by qPCR and compared to an actin control. **A)** All samples have been compared to the DMSO treated sample for the same siRNA treatment. **B)** All samples have been compared to the DMSO treated control siRNA sample. Error bars represent standard deviation with a minimum of 5 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$, and **** $p < 0.0001$.

Microprocessor knockdown also shows an increase in unspliced pri-miR-17~92a in the DMSO treated sample (**Figure 5.17B**). While unspliced pri-miR-122 may also increase following Microprocessor knockdown, the increase in pri-miR-17~92a is greater. Due to the variability across the repeats, this effect is not statistically significant, although all replicates do show an increase.

When the pla B treated cells are compared to the equivalent control or DGCR8 knockdown DMSO treated cells, only the DGCR8 knockdown pla B treated cells show a significant decrease in the levels of chromatin-associated unspliced pri-miR-23a~27a~24-2 (**Figure 5.18A**). However, the fold decrease is much smaller than that seen for unspliced pri-miR-122 and pri-miR-17~92a, and there is no significant difference between the two pla B treated samples, so it is unlikely that the Microprocessor has contributed to this effect on DGCR8 siRNA treated cells.

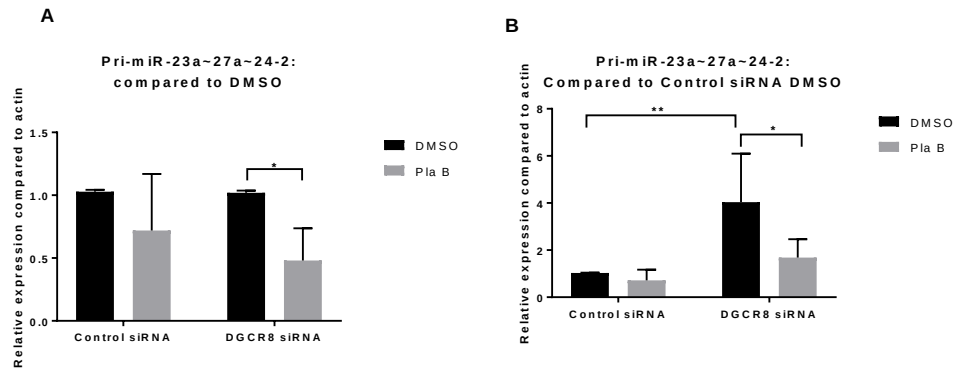


Figure 5.18: Splice inhibition using pla B and Microprocessor knockdown effects the levels of endogenous chromatin-associated unspliced pri-miR-23a~24-2~27a. A graph representing the relative expression of pri-miR-23a~24-2~27a in Huh7 cells analysed by qPCR and compared to an actin control. **A)** All samples have been compared to the DMSO treated sample for the same siRNA treatment. **B)** All samples have been compared to the DMSO treated control siRNA sample. Error bars represent standard deviation with a minimum of 5 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$ and ** $p < 0.01$.

As with chromatin-associated pri-miR-17~92a, the levels of unspliced pri-miR-23a~27a~24-2 significantly increase following DGCR8 knockdown (**Figure 5.18B**). This indicates that Microprocessing may be having a negative effect on the transcription of pri-miR-23a~27a~24-2, but it could also indicate that the lack of Microprocessing is leading to the unspliced transcript being retained in the chromatin.

As indicated in figure **5.10**, pla B has little effect on the levels of chromatin-associated unspliced pri-miR-29a~29b-1, although it may lead to a small increase, and there is no difference between the control and DGCR8 siRNA treated samples (**Figure 5.19A**). The levels of chromatin-associated unspliced pri-miR-29a~29b-1 also increase following Microprocessor knockdown compared to the control siRNA treated cells, with or without pla B treatment (**Figure 5.19B**). As this effect is also shown for unspliced pri-miR-23a~27a~24-2, pri-miR-17~92a, and to a lesser degree pri-miR-122, it indicates that Microprocessing may be having a positive effect on transcription or more probably that transcripts that have not been cleaved by Microprocessor are being retained on the chromatin.

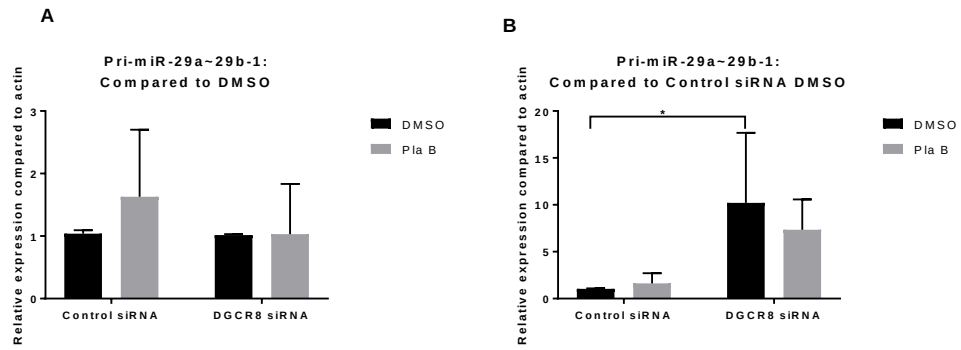


Figure 5.19: Splice inhibition using pla B has no effects on the levels of endogenous chromatin-associated unspliced pri-miR-29a~29b-1. A graph representing the relative expression of pri-miR-29a~29b-1 in Huh7 cells analysed by qPCR and compared to an actin control. **A)** All samples have been compared to the DMSO treated sample for the same siRNA treatment. **B)** All samples have been compared to the DMSO treated control siRNA sample. Error bars represent standard deviation with a minimum of 5 biological replicates and p-values calculated using a two-way ANOVA with * $p < 0.05$.

5.7: The effect of splice inhibition on pri-miR-122 in CRISPR-Cas9 knockout cells

The data shown in chapter 5.6 shows a lot of variation between samples, resulting in more error which may be hiding the significance of some of the results. This variability may be due to the amount of treatment the cells undergo. While pla B does not affect cell viability, pla B treatment in addition to Microprocessor knockdown may affect cells in a way that cannot be controlled for. To get around this issue, Huh7.5 cells which were gene edited by CRISPR-Cas9 to knock out either Drosha or the miR-122 hairpin were used (A kind gift from Charles Rice, Rockefeller University). These cells will be referred to as Δ Drosha or Δ 122 respectively. The Δ 122 cells have a 38 bp deletion in the hairpin region, and as a result lack most of the miR-122 hairpin, including the seed region. The Δ Drosha cells have a 166 bp deletion within exon 9 which abolishes Drosha protein production²⁴⁴.

To understand the effects of splice inhibition by pla B treatment on these cell lines, the cell lines were treated with 1 μ M pla B for 4 hours before chromatin-associated RNA extraction. The effect of splice inhibition was analysed by qPCR as previously used in this chapter. Huh 7.5 cells used to create the mutant cells lines were also obtained and used as a comparison in this experiment.

Wild-type Huh 7.5 cells respond in the same way to pla B treatment as Huh7 cells, and show a significant decrease in the levels of chromatin-associated spliced pri-miR-122. Similarly, $\Delta 122$ also show a significant decrease in the levels of chromatin-associated spliced pri-miR-122, as does the Δ Drosha cell line (**Figure 5.20**). This indicates that splice inhibition has the expected effect on the levels of spliced pri-miR-122.

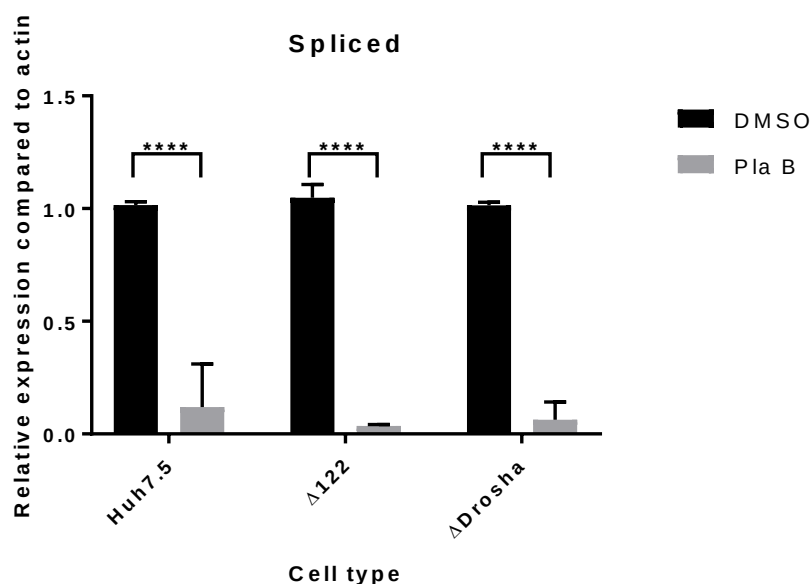


Figure 5.20: Splice inhibition using pla B affects the levels of endogenous chromatin-associated spliced pri-miR-122 in Huh 7.5 knock out cells. A graph representing the relative expression of spliced pri-miR-122 in Huh7 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated sample. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$ and **** $p < 0.001$.

As seen in Huh7 cells, the levels of unspliced pri-miR-122 decreases following pla B treatment to inhibit splicing in Huh7.5 cells (**Figure 5.21**). A similar decrease is also seen in the $\Delta 122$ cell line. This indicates that the Microprocessor is not influencing the effects of pla B on pri-miR-122 transcription.

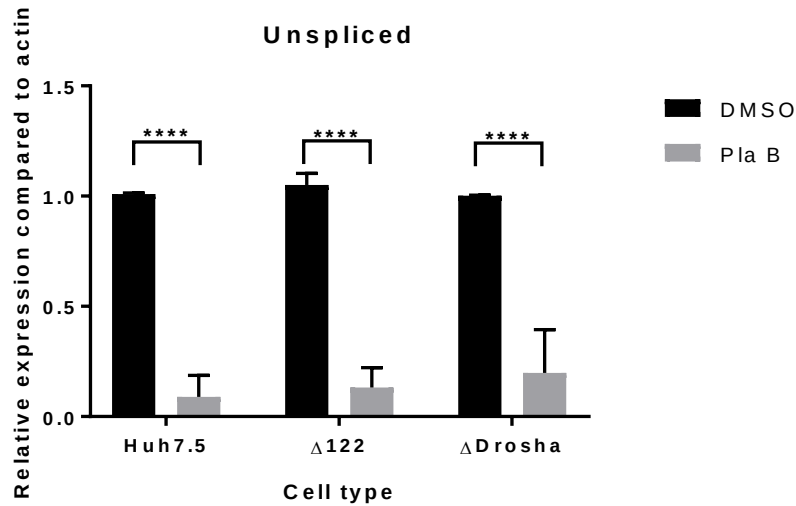


Figure 5.21: Splice inhibition using pla B affects the levels of endogenous chromatin-associated unspliced pri-miR-122 in Huh 7.5 knock out cells. A graph representing the relative expression of unspliced pri-miR-122 in Huh7 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated sample. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$ and **** $p < 0.001$.

The level of hairpin sequence retained in the pri-miR-122 transcript is shown to decrease in Huh 7.5 cells following pla B treatment (**Figure 5.22**). The decrease seen here is similar to the decrease seen in the levels of chromatin-associated unspliced pri-miR-122 following splice inhibition (**Figure 5.21**). As there is little difference between the levels of unspliced pri-miR-122 and the levels of miR-122 retained in the primary transcript, it shows that the decrease in hairpin signal is due to transcription, especially as the effect is seen in cells where the context of the hairpin has been removed in the ΔDrosha and Δ122 cells.

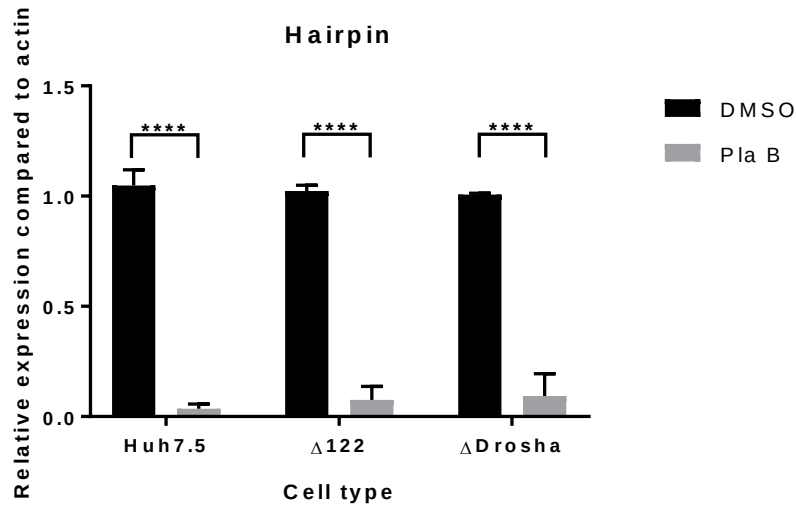


Figure 5.22: Splice inhibition using pla B affects Drosha cleavage of the miR-122 hairpin in chromatin-associated pri-miR-122 in Huh 7.5 knock out cells. A graph representing the relative expression of pri-miR-122 hairpin in Huh7.5 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated sample. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$ and **** $p < 0.001$.

As seen previously, when treated with pla B to inhibit splicing, transcriptional read through of pri-miR-122 decreases similar to the decrease in unspliced pri-miR-122. This indicates that splice inhibition is affecting pri-miR-122 at the level of transcription (**Figure 5.23**). A similar effect is seen for the $\Delta 122$ cells. Interestingly, pla B treatment has no effect on transcriptional read through in Huh7.5 cells lacking Drosha. In both the Δ Drosha and $\Delta 122$ cell lines, the hairpin is unable to be processed, so the miRNA context has been removed from the lncRNA. As the Δ Drosha cells show a different effect on transcriptional termination following splice inhibition than $\Delta 122$ cells, this could indicate that the effect seen for Δ Drosha is not a true effect. It could also indicate that the Drosha knockdown is having a secondary effect on the cells.

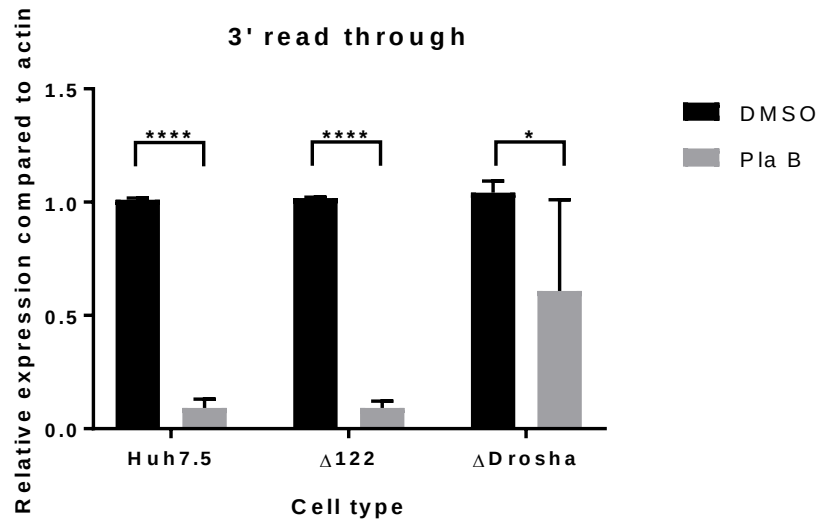


Figure 5.23: Splice inhibition using pla effects transcriptional termination of chromatin-associated pri-miR-122 in Huh 7.5 knock out cells. A graph representing the relative expression of pri-miR-122 transcriptional read through in Huh7.5 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated sample. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$ and **** $p < 0.001$.

The effect of splice inhibition on unspliced pri-miR-17~92a was also analysed in the Huh7.5 and ΔDrosha cell lines (**Figure 5.24**). In both cell lines, the levels of unspliced pri-miR-122 decrease significantly and to a similar degree. This indicates that Microprocessing has no effect on the splicing of this transcript.

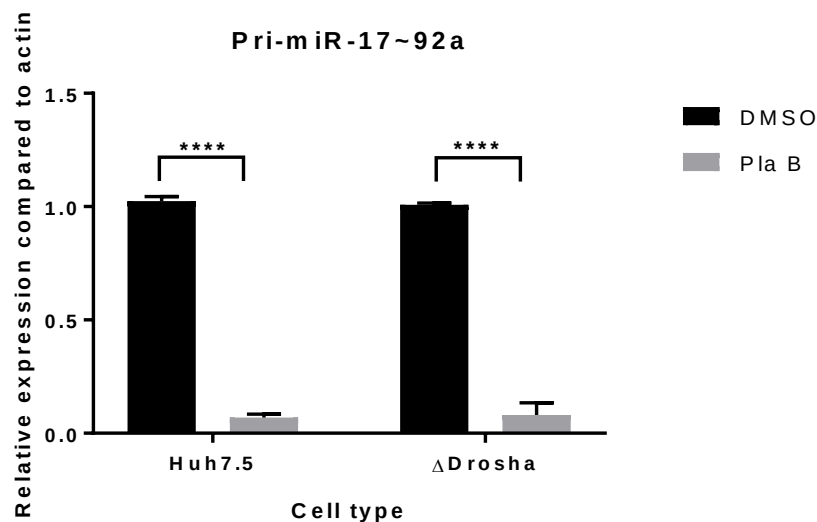


Figure 5.24: Splice inhibition using pla B effects transcriptional termination of chromatin-associated unspliced pri-miR-17~92~92a in Huh 7.5 knock out cells. A graph representing the relative expression of unspliced pri-miR-17~92a in Huh7.5 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated sample. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values

calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$ and **** $p < 0.001$

Similarly, the levels of unspliced pri-miR-23a~27a~24-2 also decreased with pla B treatment to inhibit splicing (**Figure 5.25**). This data agrees with the initial data looking at chromatin-associated pri-miR-23a~27a~24-2 following pla B treatment in Figure 5.9.

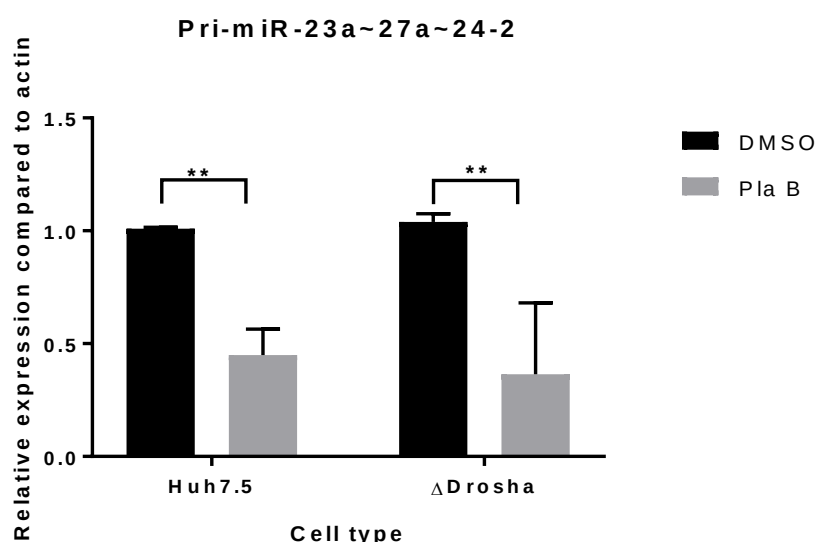


Figure 5.25: Splice inhibition using pla effects transcriptional termination of chromatin-associated unspliced pri-miR-23a~24-2~27a in Huh 7.5 knock out cells. A graph representing the relative expression of unspliced pri-miR-23a~27a~24-2 in Huh7.5 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated sample. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$ and **** $p < 0.001$

As expected, the levels of unspliced pri-miR-29a~29b-1 transcripts in these cells lines show little to no effect of splice inhibition (**Figure 5.26**). As with the other Microprocessor-terminated lnc-pri-miRNAs, the loss of Microprocessor function seems to have no effect on the levels of unspliced transcript.

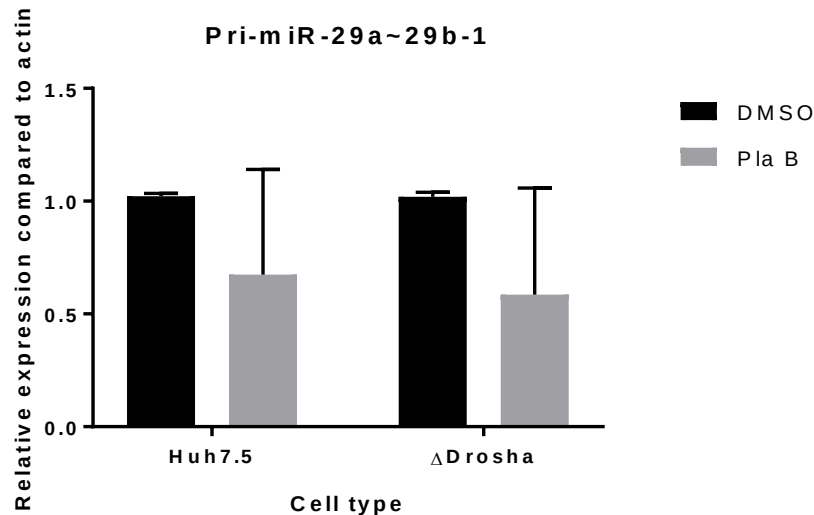


Figure 5.26: Splice inhibition using pla B effects transcriptional termination of chromatin-associated unspliced pri-miR-29a~29b-1 in Huh 7.5 knock out cells. A graph representing the relative expression of unspliced pri-miR-29a~29b-1 in Huh7.5 cells analysed by qPCR and compared to an actin control. All samples have been compared to the DMSO treated sample. Error bars represent standard deviation with a minimum of 3 biological replicates and p-values calculated using a one-way ANOVA.

5.8: Discussion

In this chapter, the effect of splice inhibition has been examined at the level of transcription using a protocol that isolates RNA from the chromatin. When treating Huh7 cells with pla B for 4 hours, the levels of chromatin-associated spliced and unspliced pri-miR-122 decrease. This agrees with the work in chapter 4, which showed the same effect with pla B but on total RNA. Therefore, splice inhibition reduces the transcription of pri-miR-122. This effect is not unique for pri-miR-122, with the transcription of pri-miR-17~92a similarly affected, and to a lesser extent, pri-miR-23a~27a~24-2.

Interestingly, Microprocessor knockdown may have a positive effect on transcription. The levels of unspliced and spliced pri-miR-122 show small increases in expression compared to control, however, there is a lot of variability within this data which makes drawing conclusions difficult. However, Microprocessor knockdown also leads to an increase in the levels of unspliced pri-miR-17~92a, pri-miR-23a~27a~24-2 and pri-miR-29a~29b-1. This could indicate that the lack of Microprocessor cleavage is leading to the transcript being retained on the chromatin.

In this chapter, as well as the previous chapter, there is some variability between repeats. As cell confluency is known to increase the expression of mature miRNA, but not the expression of the Microprocessor¹⁸⁰, chromatin-associated RNA was extracted from over-confluent Huh7 cells treated with pla B for 4 hours and compared to normally confluent cells. Confluency was found to have little effect on the levels of unspliced and spliced pri-miR-122, or unspliced pri-miR-17~92a, pri-miR-23a~27a~24-2 and pri-miR-29a~29b-1. Interestingly, confluency did have an effect on the levels of hairpin retained in the pri-miR-122 transcript following splice inhibition, with the decrease in hairpin signal following pla B treatment being significantly greater in the over-confluent cells. However, it would be expected that the levels of 3' read through would be similar to hairpin levels considering that pri-miR-122 transcription termination is Microprocessor mediated; the data suggests that the levels of transcriptional read through are the same for over confluent and normally confluent Huh7 cells treated with pla B.

Other than understanding the effect confluency has on pri-miR-122, attempts were made to reduce variability within the chromatin-associated RNA data. It is possible that DNA contamination could be present in the RNA samples. The original protocol used a short proteinase K and DNase treatment before acid-phenol chloroform precipitation, and additional chloroform precipitations had no noticeable effect. Then, a range of RNA extractions and DNase treatments were tried, and the RNA was used in a qPCR using the pri-miR-122 primers and included no RT controls. The standard RNA extraction protocol, including the DNase step and acid-phenol chloroform precipitation, removed all traces of DNA. All qPCRs using chromatin-associated RNA were analysed for DNA contamination using a no RT control using the hairpin primers.

Huh7.5 cells that had been gene-edited by CRISPR-Cas9 (as well as the original Huh7.5 cells) to delete the miR-122 hairpin (Δ 122) and Drosha (Δ Drosha) were obtained from the Rice lab²⁴⁴. The effect of pla B on chromatin-associated pri-miR-122 for these cell lines were analysed by qPCR. Huh7.5 cells are derived from the Huh7 cell line, but lack RIG-I which is involved in the innate immune response within the cytoplasm, and as both cell lines respond similarly to pla B, this indicates that RIG-I does not affect pri-miR-122 processing.

When the miRNA context of pri-miR-122 is removed in the $\Delta 122$ or Δ Drosha cell lines, the data indicates pla B has a similar effect on the levels of unspliced and spliced pri-miR-122 compared to Huh7.5 cells. The levels of miR-122 hairpin detected by qPCR to be retained in the pri-miRNA transcript also decrease similarly to Huh7 cells.

As the Microprocessor mediates transcriptional termination of pri-miR-122, it would be expected that the levels of hairpin retained in the transcript would correlate with the levels of transcriptional read through. In the Huh7.5 and $\Delta 122$ cells, the levels of hairpin and 3' read through decrease similarly following splice inhibition. However, the levels of read through for splice inhibited Δ Drosha cells do not decrease as strongly as the level of hairpin, or as strongly compared to Huh7.5 and $\Delta 122$ cells. If the effect seen in Δ Drosha is a true effect, this could indicate that transcriptional termination is inhibited. As it is known that Drosha is required to terminate pri-miR-122 transcription²⁰⁶, it would be expected that the level of transcriptional read through would increase. However, an increase in the levels of hairpin qPCR product would also be expected but was not seen. It is also possible that an unknown processing event is occurring. Alternatively, the effect could also be due to sampling error.

Chapter 6: Discussion

6.1: The effect of genomic location on miRNA biogenesis

While the majority of miRNAs are known to reside in introns, a proportion of miRNAs are known to be exonic, and while miRNAs were initially thought to be located within protein-coding genes, many miRNAs are now known to be located within a lncRNA^{95,96,99–101}. The understanding that miRNAs can be expressed from different genomic locations creates more questions. In particular, how are intronic and exonic miRNAs expressed and are they processed in the same way? Also, what effect does miRNA expression from an exon have on the host mRNA and are these processes somehow regulated?

Much of the current literature on miRNA biogenesis focuses on intronic miRNAs. It is known that miRNA biogenesis occurs co-transcriptionally^{97,109,128}. It is also known that both splicing and Microprocessing can occur on the same transcript, with Microprocessing occurring before splicing^{128,186}. It has also been shown that the Microprocessor cleavage of an intronic miRNA can lead to the surrounding introns being more rapidly spliced, as well as Microprocessor knockdown leading to an increase in splicing¹⁰². In comparison, not much is known about how miRNA expression is regulated when expressed from an exon. This is an important question, as Microprocessing is likely to lead to the destruction of the host exon, and therefore lead to a reduction in mature mRNA and protein production. Interestingly, there is an example of a mechanism that controls the expression of miRNA from an exon and this is mediated by a protein involved in splicing, KHSRP¹⁹⁷. The expression of KHSRP leads to an increase in miR-198 expression, while when the expression of KHSRP is decreased, more spliced host transcript is detected. While this is a positive indication that exonic miRNA expression is regulated with respect to the host transcript, other mechanisms have not yet been identified.

To understand how the genomic location can affect miRNA biogenesis, pre-miR-122 was inserted into different locations of a β -globin gene. The β -globin gene has frequently been used to understand co-transcriptional processing, as well as being used as a miRNA expression system^{21,128,223,238}. β -globin plasmids containing pre-miR-122 were

transfected into HeLa cells which do not express β -globin or miR-122. Interestingly, this work showed that when pre-miR-122 was located within an exon or an intron of a protein coding gene, the levels of mature miR-122 increased in comparison to miR-122 expressed from a lncRNA. Surprisingly, the levels of mature miR-122 expressed were similar from the intron and exons. This suggests genomic location does have an effect on mature miRNA expression, however not necessarily in the way that was expected. Interestingly, the expression of miR-122 from a lncRNA was lower than from an exon. This could indicate that the Microprocessor has less time to process a hairpin as miR-122 is located at the 3' end of the lncRNA transcript. The expression of pre-miR-122 from the 3' region of the β -globin transcript is reduced compared to the lncRNA context. This would indicate that not only can the genomic context of miRNA location effect expression of mature miRNA, but overall position in the transcript too, with miRNA located more towards the 3' end of the transcript less efficiently expressed.

To understand if 3' to 5' positioning, along with genomic context, has an effect on miRNA expression, the β -globin plasmid could be used. By expressing pre-miR-122 in more locations, such as intron 1 and exon 1, or pre-miR-122 expression from different insertion points within a genomic location, such as 3', middle or 5' of intron 1, the effect of 3' to 5' positioning with respect to genomic context could be analysed by qPCR (**Figure 6.1**).

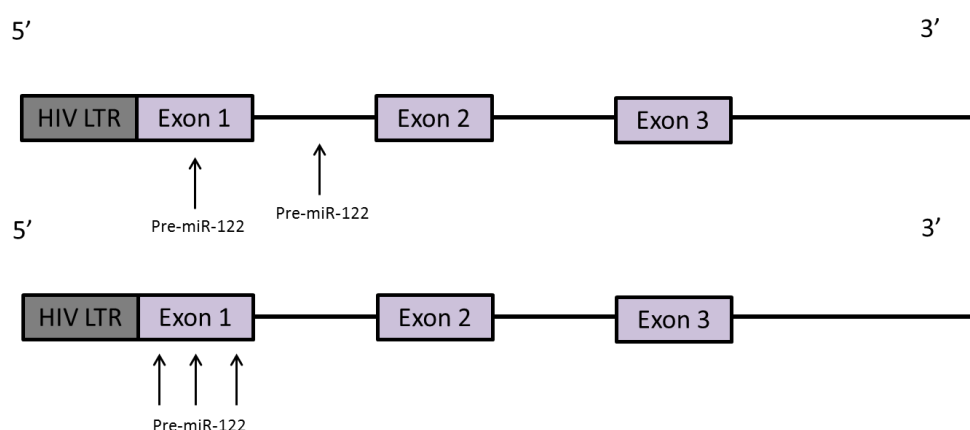


Figure 6.1: Schematic of miR-122 β -globin constructs. **A)** β -globin reporter under the control of a HIV LTR promoter showing the possible locations of pre-miR-122 in exon 1 and intron 1. **B)** Possible pre-miR-122 in exon 1.

Exonic miRNA expression would lead to cleavage within the exon which should lead to the loss of the host mRNA. If Microprocessing occurred, this would lead to transcription being prematurely terminated, however, it is possible for post-transcriptional Microprocessing to occur, which would lead to degradation of the host mRNA. Interestingly, when pre-miR-122 was expressed from an exon, a similar level of mature miR-122 was expressed compared to an intronic location, and spliced β -globin levels were only slightly reduced compared to the empty vector control. While it is known that intronic miRNAs are cleaved before splicing¹⁰² it is possible that that Microprocessing is also occurring faster than splicing when the miRNA is exonic. It is also possible that miRNA cleavage from an exon could result in exon skipping, and that this was detected to a small degree on the northern blot; exon skipping is unlikely to be a minor effect. Exon skipping of exon 2 or exon 3 should not affect the qPCR data as the primers span the exon 2- exon 3 junction, however, primers spanning the exon 1-exon 3 junction would be able to detect exon skipping when pre-miR-122 is expressed from exon 2.

To understand whether the spliceosome is involved in this regulation, HeLa cells transfected with a β -globin plasmid could be treated with a splice inhibitor. Attempts were made to do this with the splice inhibitor pla B. Unfortunately, this data was very inconsistent and has not been shown. While inconsistent data could indicate that splice inhibition is having an effect on miR-122 or β -globin expression, it is more likely that pla B treatment following a plasmid transfection was negatively affecting the cells. 4 hours of pla B has been shown to have no cytotoxic effect on HeLa cells, but this did not include a plasmid transfection. It was also noted that after pla B treatment, there was visible cell death in the transfected cells. Similarly, splice site β -globin mutants, abolishing either the 5' SS, the 3' SS or both, could be used. This would reduce the amount of treatment the cells undergo and be less likely to lead to cell death. It would also only affect the gene construct, unlike pla B treatment which globally inhibits splicing and may lead to secondary effects which could affect the results.

It is also possible that the relationship between miRNA expression and splicing from the different β -globin plasmids may be due to the pre-miR-122 hairpin and the flanking sequence and not the location within the gene. To test this, another miRNA could be inserted within the β -globin

gene (**Figure 6.2**). Attempts were made to do this with miR-155, which is also not expressed in HeLa cells and is located within a lncRNA, although amplifying pre-miR-155 from genomic DNA proved unsuccessful.

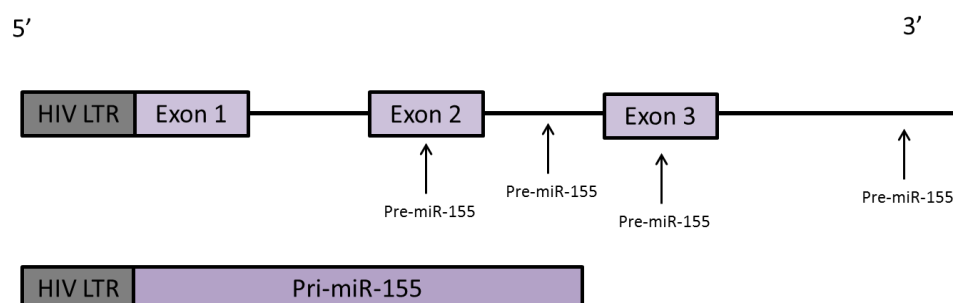


Figure 6.2: Schematic of miR-155 reporter constructs. A) β -globin reporter under the control of a HIV LTR promoter showing the possible locations of wild-type and Drosha-cleavage impaired pre-miR-155. **B)** Full length pri-miR-155 lncRNA under the same HIV LTR reporter.

This work could be elaborated further to look at the competition between two miRNA expressed from different genomic locations of the same transcript. For example, pre-miR-122 could be expressed from intron 2, whereas pre-miR-155 might be expressed from exon 2 (**Figure 6.3**). qPCR analysis of the mature miRNA expressed from this construct would provide information about how the expression of two miRNA from one construct is regulated and whether intronic or exonic miRNA predominates when both are present.

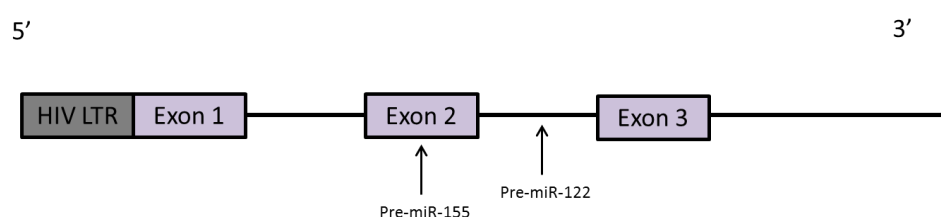


Figure 6.3: Schematic of miR-155 and miR-122 within the same reporter constructs. A) β -globin reporter under the control of a HIV LTR promoter showing the possible locations of wild-type pre-miR-122 and pre-miR-155.

During these experiments, the levels of β -globin protein expressed was not examined by western blot. This was because the work focused on the nuclear processing events and it is reasonable to assume that the spliced mRNA detected by qPCR and northern blot would be exported to the cytoplasm. It would be interesting to know what effect Microprocessing would have downstream on transcription. Western blotting could be used

to measure the levels of β -globin protein when pre-miR-122 is expressed from an intron or exon. Polysome profiling could also be used as an indication to whether the mRNA is being associated to ribosomes would give an indication on how much functional mRNA is produced should western blot not be successful.

6.2: Splice inhibition reduces transcription efficiency for endogenous pri-miR-122

The other aim of this thesis was to understand how the genomic context of a miRNA can affect its expression and how splicing influences miRNA biogenesis from an exon of a lncRNA. Little is known about how miRNAs are expressed from a lncRNA or an exon, and the medically relevant miR-122 is located within the terminal exon of a lncRNA. Interestingly, there is evidence that RNA processing can occur differently for lncRNA compared to protein-coding RNA; a novel mechanism of transcriptional termination was discovered that is unique to lnc-pri-miRNAs, with a subset of lnc-pri-miRNA such as pri-miR-122 using the Microprocessor to terminate transcription²⁰⁶. Understanding how the biogenesis of miR-122 is regulated at a transcriptional level may lead to a better understanding of how it is so highly expressed and liver specific and may lead to new drug discoveries.

To measure the effect of splice inhibition, Huh7 cells which endogenously express pri-miR-122 were treated with pla B. Similarly, HeLa cells ectopically expressing pri-miR-122 were also treated with pla B. qPCR analysis of the levels of spliced and unspliced pri-miR-122 showed that endogenously expressed, and overexpressed, pri-miR-122 was affected by pla B differently (**Figure 6.3**). The ectopically expressed pri-miR-122 behaved as expected; spliced decreased, unspliced increased. While the levels of endogenous spliced pri-miR-122 decreased, the levels of unspliced transcript also decreased.

Table 6: The effect of splice inhibition of pri-miR-122

	Overexpressed pri-miR-122 in HeLa	Endogenous pri-miR-122 in Huh7
Spliced		
Unspliced		

Both spliced and unspliced pri-miR-122 levels decreasing would suggest that splice inhibition with pla B is having a negative effect on transcription (**Table 6**). This same effect on transcription was also seen with two other drug inhibitors, spliceostatin A and meayamycin which both act on the same pathway as pla B (unpublished work by Catherine Jopling) to inhibit splicing at the level of SF3b and are generally known not to have an effect on transcription⁷¹. To confirm that the effect of splice inhibition is not unique to the drug inhibitors, siRNA targeting the splicing factor SF3b1 and morpholinos targeting the U1 and U2 snRNPs were used which also showed a similar effect on transcription.

To understand this further, RNA in the process of being transcribed was extracted from Huh7 cells treated with pla B and analysed by qPCR. This also showed the same decrease in the levels of spliced and unspliced pri-miR-122 transcripts. As pri-miR-122 transcription is terminated by the Microprocessor the levels of transcriptional read through and the levels of hairpin retained in the transcript were measured by qPCR. While transcriptional read through and the level of hairpin decreased similarly with pla B, they were increased relative to the level of unspliced transcript. This would indicate that splice inhibition is affecting Microprocessing. However, this result was not substantiated in later experiments, indicating that the initial effect on Microprocessing may not be a true effect. While the effect of confluency on pri-miR-122 processing was evaluated, which showed that splice inhibition did affect Microprocessing at high confluence as indicated in the literature¹⁸⁰ it is possible that subtle differences in cell confluence between experimental repeats affected the data. It is also possible that the

concentration of pla B varied between repeats. Each batch of pla B was tested against the previous batch to ensure consistency, and 10 mM aliquots of pla B were stored prior to being diluted for use in cell culture. While attempts were made to keep the pla B consistent, small changes could cause differences between repeated experiments, especially in conjunction with small changes in cell confluency. To minimise the small difference that might be caused by pla B concentration, 1mM stocks could be used instead and this would remove any dilution errors.

The expression of endogenous mature miR-122 from Huh7 and ectopically expressed mature miR-122 from Huh7 was also analysed by qPCR. This showed that 4 hours of pla B treatment had little effect on the levels of mature miR-122. This is not surprising considering miR-122 is very stable²⁴⁶. To try and understand the effect splicing inhibition has on mature miR-122 expression, a different method of looking at newly transcribed RNA could be used, such as 4-SU.

4-SU labelling involves treating cells with 4-thioluridine which is incorporated into newly transcribed RNA which can then be biotinylated and separated from the total RNA pool²⁴⁷. This would allow newly synthesised mature miR-122 to be isolated and detected by qPCR. Attempts were made to use 4-SU labelling to look at the effects of endogenous pri-miR-122 initially as these are larger transcripts and would be easier to detect. Unfortunately, the background levels of unlabelled pri-miR-122 were too high to draw any relevant conclusion from these experiments. Attempts were also made to use a commercially available kit to label newly transcribed RNA which uses a different uracil analogue, ethylene uracil. Again, there were high levels of background contamination, and this protocol would require further optimisation.

6.3: The efficiency of transcription for lnc-pri-miRNAs

It has been shown elsewhere that pri-miR-122 transcription termination is mediated by the Microprocessor and not by the canonical cleavage and polyadenylation pathway with DGCR8 knockdown leading to an increase in transcriptional read through²⁰⁶. This work also showed that Microprocessor-terminated transcription is not unique to pri-miR-122, but occurs on a variety of lncRNA containing miRNAs but not protein coding pri-miRNA

genes. DGCR8 knockdown was shown to lead to an increase in transcriptional read through for both lnc-pri-miR-17~92a and lnc-pri-miR-29a~29b-1a while lnc-pri-miR-23a~27a~24-2 did not show a significant increase.

qPCR analysis was used to detect the levels of unspliced transcripts for these three lnc-pri-miRNAs. This showed that unspliced pri-miR-17~92a decreased significantly following splice inhibition, but no effect was seen for unspliced pri-miR-29a~29b-1. This indicates that splicing is required for the efficient transcription for a subset of lnc-pri-miRNAs.

Splice inhibition had little to no effect on the levels of unspliced pri-miR-23a~27a~24-2. Table 7 shows the effect splice inhibition had on the unspliced pri-miRNA transcript compared to the effect on transcriptional read through when DGCR8 is knocked down in Huh7 as described elsewhere²⁰⁶.

Table 7: The effect of Microprocessor knockdown or splice inhibition of lnc-pri-miRNAs

	Transcriptional read through with DGCR8 KD	Effect of splice inhibition on unspliced pri-miRNA
Pri-miR-122	X 2.2 	
Pri-miR-17	X 4 	
Pri-miR-23	X 1.8 	 or 
Pri-miR-29	X 23 	 or 

Taken together, this could indicate that splice inhibition may have an effect on a subset of lnc-pri-miRNAs, or even a subset of Microprocessor terminated lnc-pri-miRNAs and shows that the transcription of lnc-pri-miRNAs are more complex than is currently understood.

To understand the effect other processing factors might be having on lnc-pri-miRNA expression, knockdown experiments using siRNA to target other co-transcriptional processes, such as capping could be used. Another experiment that could be used to assess this would be Chromatin immunoprecipitation (ChIP). This would involve crosslinking the RNA and the proteins associated with them, and using an antibody targeting a co-transcriptional protein to pull down the associated proteins and RNA. While qPCR could be used to look at the specific lnc-pri-miRNAs, a transcriptomics approach will allow a global view on how lnc-pri-miRNAs are processed. This could also be combined with knockdown experiments to elucidate which co-transcriptional processes are important for the transcription of specific lnc-pri-miRNAs.

Attempts were made to look at the effect of splice inhibition on the Drosha cleavage of the five miRNAs within lnc-pri-miR-17~92a. While these miRNAs form a cluster and are all transcribed at the same time, it has been shown that the levels of mature miRNA expression from the transcript is not equal due to the tertiary structure produced by the hairpins²¹⁹. qPCR analysis using primers spanning several Drosha cleavage sites in this miRNA cluster were inconsistent and it would be necessary to optimise qPCR primers and conditions to understand how Microprocessing of this transcript is affected.

This work has only looked at four lncRNAs which contain miRNAs. To understand how lncRNA transcription is regulated, a transcriptomics approach is necessary. This initially could be a continuation of the work discussed here using pla B to inhibit splicing. As pla B has been shown to reduce transcription efficiency, this could indicate which lnc-pri-miRNAs require the association of the spliceosome during transcription, with mature miRNA expression also analysed. Next-generation sequencing (NGS) of chromatin-associated RNA following pla B treatment to inhibit splicing compared to control cells would allow for direct analysis of the effects on transcription. Additionally, the sequence reads across Drosha cleavage sites could be analysed to measure the effect on Microprocessing. While the work undertaken here has not shown much effect of Microprocessor knockdown on the transcription of four lnc-pri-miRNAs, it is possible that Microprocessor knockdown may affect other lnc-pri-miRNAs,

and NGS following Microprocessor knockdown with or without splice inhibition may be informative.

6.4: Pri-miR-122 is rapidly degraded

In chapter 4, the effect of Microprocessor knockdown and splice inhibition on endogenous total pri-miR-122 was assessed by qPCR. As the data suggested that splice inhibition was having a negative effect on transcription, the experiment was repeated to look at the RNA during transcription. As these experiments produce a lot of data, it is simpler to look at this data as a table and table 8 shows the expression of total and chromatin-associated pri-miR-122 transcripts relative to the DMSO control.

Table 8: The relative expression of total and chromatin-associated pri-miR-122 transcripts following splice inhibition and Microprocessor knockdown.

		Spliced		Unspliced		Hairpin		3' read through	
		Total	Chromatin	Total	Chromatin	Total	Chromatin	Total	Chromatin
No siRNA	DMSO	1	1	1	1	1	1	1	1
	Pla B	0.04	0.04	0.15	0.11	0.41	0.41	0.36	0.36
Control siRNA	DMSO	1	1	1	1	1	1	1	1
	Pla B	0.07	0.07	0.09	0.17	0.45	0.15	0.17	0.11
DGCR8 siRNA	DMSO	1	1	1	1	1	1	1	1
	Pla B	0.02	0.02	0.42	0.24	0.33	0.11	0.11	0.19

When looking at the levels of spliced pri-miR-122, it is very clear that there is no difference between the total and the chromatin-associated transcripts following pla B treatment to inhibit splicing, and the same is true for the 3' read through. This indicates that once released from the chromatin, these transcripts are rapidly degraded. However, when the cells were treated with DGCR8 siRNA and pla B, the levels of unspliced pri-miR-122 are higher in the total than the chromatin. This could indicate that the lack of Microprocessing is protecting the transcript from decay. Similarly, the levels of hairpin transcripts are also enriched in the total compared to the chromatin. As these transcripts show a lack of Microprocessing, this indicates that Microprocessor cleavage leads to decay.

6.5: The co-transcriptional mRNA processing factors and their interactions

The current literature has clearly shown that Microprocessing occurs during transcription. It has also been shown that the Microprocessor protein Drosha, as well as other RNA processing factors associates with the CTD of RNA polymerase II during transcription. Some factors, such as CBP80 have been shown to have a role in splicing and also interact with the Microprocessor¹¹⁴. The complexity of Microprocessor regulation has become apparent in recent years, with many cofactors being identified.

It is possible that an unknown factor, or even multiple factors, may interact with the mRNA processing machinery during transcription and it is already known that co-transcriptional processing of lncRNAs occur differently to protein-coding genes⁸⁹. In this thesis, it has been shown that inhibiting splicing has an effect on the transcription of pri-miR-122 across the transcript, and it is possible that a similar effect occurs for pri-miR-17~92a. However, it is unlikely that splicing has an effect on Microprocessor efficiency for these lnc-pri-miRNAs. To further elucidate a relationship between the transcription of lnc-pri-miRNAs and co-transcriptional processes, RNA immunoprecipitation (RIP) could be used to look at the interaction of the Microprocessor proteins DGCR8 and Drosha with the pri-miRNAs in combination with splice inhibition or the knockdown of other co-transcriptional processes. Coupled with NGS, this could provide a lot of new information about how miRNA biogenesis can be regulated when expressed from a lncRNA.

As previously mentioned, ChIP could be used to look at the interaction with RNA polymerase II and lnc-pri-miRNA transcripts. Different antibodies recognising different phosphorylated forms of the CTD could also be used, and CTD phosphorylation mutants could also be used to understand how the interaction of the CTD and proteins involved in transcription may regulate miRNA expression from a lncRNA. Similarly, ChIP could be used to analyse the histone marks surrounding lnc-pri-miRNAs to understand how epigenetic regulation could alter miRNA biogenesis. Techniques such as NET-seq, which do not use crosslinking agents which may affect chromatin structure would allow for detailed analysis of the nascent RNA and the CTD code⁷².

While the majority of this thesis is looking at lnc-pri-miRNAs, and how splicing is necessary for the transcription of a subset of these, how miRNAs are expressed from an exon of a protein-coding gene is still not well understood. With NGS experiments, exonic miRNAs could also be analysed, and if compared with exonic miRNAs expressed from a lncRNA, this may provide more information about how transcription of protein-coding genes differ from lncRNA.

6.6: Final remarks

The underlying aims of this thesis was to elucidate how the genomic location of a miRNA may affect its location and to understand how this expression may be regulated with respect to splicing. The work undertaken has shown that a miRNA located from an exon can be efficiently expressed, and that spliced mRNA can also be strongly expressed. It has also been shown that splicing, or the association of the spliceosome is necessary for the efficient for the transcription of some lnc-pri-miRNAs, but not all. While it is already known that lncRNA transcription is different compared to protein-coding genes⁸⁹, this indicates that there may be differences in how subsets of pri-miRNA are transcribed and that the transcription of lnc-pri-miRNA may be more complex than initially thought. Understanding how miRNAs are expressed from a lncRNA, and how lnc-pri-miRNAs are transcribed in general will lead to a better understanding of the role both lncRNAs and miRNAs play in diseases, and could ultimately lead to the development of new therapeutics.

References

1. Crick F. (1958). On Protein Synthesis. The Symposia of the Society for Experimental Biology:138–166.
2. Baltimore D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226(5252):1209–1211.
3. Temin H.M., & Mizutani S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226:1211–1213.
4. Uzawa T., Yamagishi A., & Oshima T. (2002). Polypeptide Synthesis Directed by DNA as a Messenger in Cell-Free Polypeptide Synthesis by Extreme Thermophiles, *Thermus thermophilus* HB27 and *Sulfolobus tokodaii* Strain 7. *Journal of Biochemistry* 131(6):849–853.
5. Koonin E. V (2012). Does the central dogma still stand? *Biology direct* 7:27.
6. Allfrey V.G., Faulkner R., & Mirsky A.E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of RNA Synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 51(5):786–94.
7. Tan M., Luo H., Lee S., Jin F., Yang J.S., Montellier E., Buchou T., Cheng Z., Rousseaux S., Rajagopal N., et al. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146(6):1016–28.
8. Sainsbury S., Bernecky C., & Cramer P. (2015). Structural basis of transcription initiation by RNA polymerase II. *Molecular Cell Biology* 16:129–143.
9. Grünberg S., & Hahn S. (2013). Structural insights into transcription initiation by RNA polymerase II. *Trends in Biochemical Sciences* 38(12):603–611.
10. Roeder R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends in Biochemical Sciences* 21(9):327–335.
11. Heidemann M., Hintermair C., Vo?? K., & Eick D. (2013). Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1829(1):55–62.
12. Kelly W.G., Dahmus M.E., & Hart G.W. (1993). RNA polymerase II is a glycoprotein: Modification of the COOH-terminal domain by O-

- GlcNAc. *Journal of Biological Chemistry* 268(14):10416–10424.
13. Kubicek K., Cerna H., Holub P., Dev G., Pasulka J., Hrossova D., Loehr F., Hofr C., & Vanacova S. (2012). Serine phosphorylation and proline isomerization in RNAP II CTD control recruitment of Nrd1
Email alerting service Serine phosphorylation and proline isomerization in RNAP II CTD control recruitment of Nrd1. *Genes and Development* 26:1891–1896.
 14. Sims III R.J., Rojas L.A., Beck D., Bonasio R., Schüller R., Drury III W.J., Eick D., & Reinberg D. (2011). The C-Terminal DSimsomain of RNA Polymerase II Is Modified by Site-Specific Methylation. *Science* 332(April):99–103.
 15. Li H., Zhang Z., Wang B., Zhang J., Zhao Y., & Jin Y. (2007). Wwp2-mediated ubiquitination of the RNA polymerase II large subunit in mouse embryonic pluripotent stem cells. *Molecular and cellular biology* 27(15):5296–305.
 16. Zhang D.W., Rodríguez-Molina J.B., Tietjen J.R., Nemec C.M., & Ansari A.Z. (2012). Emerging Views on the CTD Code. *Genetics research international* 2012:347214.
 17. Cramer P., Bushnell D. a, & Kornberg R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science (New York, NY)* 292(5523):1863–1876.
 18. McCracken S., Fong N., Yankulov K., Ballantyne S., Pan G., Greenblatt J., Patterson S.D., Wickens M., & Bentley D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* 385(6614):357–361.
 19. Fong N., & Bentley D.L. (2001). Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: Different functions for different segments of the CTD. *Genes and Development* 15(14):1783–1795.
 20. Colgan D.F., & Manley J.L. (1997). Mechanism and regulation of mRNA polyadenylation. *Genes & development* 11(21):2755–66.
 21. Proudfoot N.J., & Brownlee G.G. (1976). 3' non-coding region sequences in eukaryotic messenger RNA. *Nature* 263(5574):211–214.
 22. Mandel C.R., Bai Y., & Tong L. (2008). Protein factors in pre-mRNA 3'-end processing. *Cellular and Molecular Life Sciences* 65(7–8):1099–1122.
 23. Wilusz J., Shenk T., Takagaki Y., & Manley J.L. (1990). A

- multicomponent complex is required for the AAUAAA-dependent cross-linking of a 64-kilodalton protein to polyadenylation substrates. *Molecular and Cellular Biology* 10(3):1244–1248.
24. Takagaki Y., Manley J.L., MacDonald C.C., Wilusz J., & Shenk T. (1990). A multisubunit factor, CstF, is required for polyadenylation of mammalian pre-mRNAs. *Genes and Development* 4(12):2112–2120.
 25. Wilusz J., & Shenk T. (1988). A 64 kd nuclear protein binds to RNA segments that include the AAUAAA polyadenylation motif. *Cell* 52(2):221–228.
 26. Takagaki Y., Ryner L.C., & Manley J.L. (1989). Four factors are required for 3' -end cleavage of pre-mRNAs. *Genes & development* 3(11):1711–1724.
 27. Rügsegger U., Beyer K., & Keller W. (1996). Purification and characterization of human cleavage factor Im involved in the 3' end processing of messenger RNA precursors. *Journal of Biological Chemistry* 271(11):6107–6113.
 28. Brown K.M., & Gilmartin G.M. (2003). A Mechanism for the Regulation of Pre-mRNA 3'?? Processing by Human Cleavage Factor Im. *Molecular Cell* 12(6):1467–1476.
 29. Terns M.P., & Jacob S.T. (1989). Role of poly(A) polymerase in the cleavage and polyadenylation of mRNA precursor. *Molecular and cellular biology* 9(4):1435–44.
 30. Kühn U., Gündel M., Knoth A., Kerwitz Y., Rüdél S., & Wahle E. (2009). Poly(A) tail length is controlled by the nuclear Poly(A)-binding protein regulating the interaction between Poly(A) polymerase and the cleavage and polyadenylation specificity factor. *Journal of Biological Chemistry* 284(34):22803–22814.
 31. Rappsilber J., Ryder U., Lamond A.I., & Mann M. (2003). Large-scale proteomic analysis of the human spliceosome. *Genome Research* 13(1):1231–1245.
 32. Herzel L., Ottoz D.S.M., Alpert T., & Neugebauer K.M. (2017). Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nature Reviews Molecular Cell Biology*. doi:10.1038/nrm.2017.63.
 33. Bieberstein N.I., Oesterreich F.C., Straube K., & Neugebauer K.M. (2012). First exon length controls active chromatin signatures and transcription. *Cell Reports* 2(1):62–68.

34. Mount S.M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Research* 10(2):459–472.
35. Lim L.P., & Burge C.B. (2001). A computational analysis of sequence features involved in recognition of short introns. *Proceedings of the National Academy of Sciences* 98(20):11193–11198.
36. Mercer T.R., Clark M.B., Andersen S.B., Brunck M.E., Haerty W., Taft R.J., Nielsen L.K., Dinger M.E., & Mattick J.S. (2010). Genome-wide discovery of human splicing branchpoints. *Genome Res* 25:290–303.
37. Taggart A.J., Lin C.L., Shrestha B., Heintzelman C., Kim S., & Fairbrother W.G. (2017). Large-scale analysis of branchpoint usage across species and cell lines. *Genome Research* 27(4):639–649.
38. Hoang T., Nguyen D., Galej W.P., Bai X., Savva C.G., Andrew J., Scheres S.H.W., & Nagai K. (2015). Europe PMC Funders Group The architecture of the spliceosomal U4 / U6 . U5 tri-snRNP. *Nature* 523(7558):47–52.
39. Rigo N., Sun C., Fabrizio P., Kastner B., & Lührmann R. (2015). Protein localisation by electron microscopy reveals the architecture of the yeast spliceosomal B complex. *The EMBO journal* 34(24):3059–3073.
40. Hoang T., Nguyen D., Galej W.P., Bai X., Oubridge C., & Andrew J. (2016). Europe PMC Funders Group CryoEM structure of the yeast U4 / U6 . U5 tri-snRNP at 3 . 7 Å resolution. *Nature* 530(7590):298–302.
41. Fabrizio P., Dannenberg J., Dube P., Kastner B., Stark H., Urlaub H., & Lührmann R. (2009). The Evolutionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome. *Molecular Cell* 36(4):593–608.
42. Schneider C., Agafonov D.E., Schmitzová J., Hartmuth K., Fabrizio P., & Lührmann R. (2015). Dynamic Contacts of U2, RES, Cwc25, Prp8 and Prp45 Proteins with the Pre-mRNA Branch-Site and 3' Splice Site during Catalytic Activation and Step 1 Catalysis in Yeast Spliceosomes. *PLoS Genetics* 11(9):1–27.
43. Liu H.-L., & Cheng S.-C. (2012). The Interaction of Prp2 with a Defined Region of the Intron Is Required for the First Splicing Reaction. *Molecular and Cellular Biology* 32(24):5056–5066.
44. Rauhut R., Fabrizio P., Dybkov O., Hartmuth K., Pena V., Chari A., Kumar V., Lee C.-T., Urlaub H., Kastner B., et al. (2016). Molecular

- architecture of the *Saccharomyces cerevisiae* activated spliceosome. *Science* 353(6306):1399–1405.
45. Galej W.P., Wilkinson M.E., Fica S.M., Oubridge C., Newman A.J., & Nagai K. (2016). Cryo-EM structure of the spliceosome immediately after branching. *Nature* 537(7619):197–201.
 46. Wan R., Yan C., Bai R., Huang G., & Shi Y. (2016). Structure of a yeast catalytic step I spliceosome at 3.4 Å resolution. *Science* 353(6302):895–904.
 47. Ohrt T., Odenwälder P., Dannenberg J., Prior M., Warkocki Z., Schmitzová J., Karaduman R., Gregor I., Enderlein J., Fabrizio P., et al. (2013). Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system. *RNA (New York, NY)* 19(7):902–15.
 48. Schwer B. (2008). A Conformational Rearrangement in the Spliceosome Sets the Stage for Prp22-Dependent mRNA Release. *Molecular Cell* 30(6):743–754.
 49. Smith C.W., Chu T.T., & Nadal-Ginard B. (1993). Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns. *Molecular and cellular biology* 13(8):4939–52.
 50. Gahura O., Hammann C., Valentová A., Puta F., & Folk P. (2011). Secondary structure is required for 3' splice site recognition in yeast. *Nucleic Acids Research* 39(22):9759–9767.
 51. Meyer M., Plass M., Pérez-Valle J., Eyraas E., & Vilardell J. (2011). Deciphering 3'ss Selection in the Yeast Genome Reveals an RNA Thermosensor that Mediates Alternative Splicing. *Molecular Cell* 43(6):1033–1039.
 52. Will C.L., & Lührmann R. (2011). Spliceosome structure and function. *Cold Spring Harbor perspectives in biology* 3(7). doi:10.1101/cshperspect.a003707.
 53. Fourmann J.B., Schmitzová J., Christian H., Urlaub H., Ficner R., Boon K.L., Fabrizio P., & Lührmann R. (2013). Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes and Development* 27(4):413–428.
 54. Turunen J.J., Niemelä E.H., Verma B., & Frilander M.J. (2013). The significant other: Splicing by the minor spliceosome. *Wiley Interdisciplinary Reviews: RNA* 4(1):61–76.
 55. Furuichi Y., & Miura K. (1975). A blocked structure at the 5' terminus

of mRNA from cytoplasmic polyhedrosis virus. *Nature* 253(5490):374–375.

56. Filipowicz W., Furuichi Y., Sierra J.M., Muthukrishnan S., Shatkin A.J., & Ochoa S. (1976). A protein binding the methylated 5'-terminal sequence, m⁷GpppN, of eukaryotic messenger RNA. *Proceedings of the National Academy of Sciences of the United States of America* 73(5):1559–63.
57. Banerjee A.K. (1980). 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiological Reviews* 44(2):175–205.
58. Shatkin A.J., Manning R.F., Gage L.P., Reichmann M.E., Furuichi Y., Morgan M., & Shatkin A.J. (1976). Capping of eucaryotic mRNAs. *Cell* 9(4 PT 2):645–53.
59. Izaurralde E., Lewis J., McGuigan C., Jankowska M., Darzynkiewicz E., & Mattaj L.W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* 78(4):657–668.
60. Lewis J.D., Izaurralde E., Jarmolowski A., McGuigan C., & Mattaj L.W. (1996). A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes and Development* 10(13):1683–1698.
61. full-text.
62. Izaurralde E., Stepinski J., Darzynkiewicz E., & Mattaj I.W. (1992). A cap binding protein that may mediate nuclear export of RNA polymerase II-transcribed RNAs. *Journal of Cell Biology* 118(6):1287–1295.
63. Flaherty S.M., Fortes P., Izaurralde E., Mattaj I.W., & Gilmartin G.M. (1997). Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proceedings of the National Academy of Sciences of the United States of America* 94(October):11893–11898.
64. Jackson R.J., Hellen C.U.T., & Pestova T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature reviews Molecular cell biology* 11(2):113–127.
65. Otto G. (2002). Ribosomal proteins mediate the hepatitis C virus IRES-HeLa 40S interaction. *RNA (New York, NY)* 8(7):913–923.
66. Lancaster A.M., Jan E., & Sarnow P. (2006). Initiation factor-independent translation mediated by the hepatitis C virus internal ribosome entry site. *RNA (New York, NY)* 12(5):894–902.
67. Kieft J.S. (2008). Viral IRES RNA structures and ribosome

- interactions. *Trends in Biochemical Sciences* 33(6):274–283.
68. Garneau N.L., Wilusz J., & Wilusz C.J. (2007). The highways and byways of mRNA decay. *Nature reviews Molecular cell biology* 8(2):113–126.
 69. Kilchert C., Wittmann S., & Vasiljeva L. (2016). The regulation and functions of the nuclear RNA exosome complex. *Nature Reviews Molecular Cell Biology* 17(4):227–239.
 70. West S., Gromak N., & Proudfoot N.J. (2004). Human 5' → 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature* 432(7016):522–525.
 71. Davidson L., Kerr A., & West S. (2012). Co-transcriptional degradation of aberrant pre-mRNA by Xrn2. *The EMBO Journal* 31(11):2566–2578.
 72. Nojima T., Gomes T., Grosso A.R.F., Kimura H., Dye M.J., Dhir S., Carmo-Fonseca M., & Proudfoot N.J. (2015). Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell* 161(3):526–540.
 73. Lowell J.E., & Sachs A.B. (1992). 3'-UTR-dependent deadenylation by the yeast poly(A) nuclease. *Genes & Development*:2088–2099.
 74. Wahle E., & Winkler G.S. (2013). RNA decay machines: Deadenylation by the Ccr4–Not and Pan2–Pan3 complexes. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1829(6):561–570.
 75. Yamashita A., Chang T.-C., Yamashita Y., Zhu W., Zhong Z., Chen C.-Y.A., & Shyu A.-B. (2005). Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nature structural & molecular biology* 12(12):1054–1063.
 76. Chen J., Chiang Y.-C., & Denis C.L. (2002). CCR4, a 3'-5' poly(A) RNA and ssDNA exonuclease, is the catalytic component of the cytoplasmic deadenylase. *The EMBO journal* 21(6):1414–26.
 77. Mittal S., Aslam A., Doidge R., Medica R., & Winkler G.S. (2011). The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4-Not complex contribute to the prevention of cell death and senescence. *Molecular biology of the cell* 22(6):748–58.
 78. Doidge R., Mittal S., Aslam A., & Winkler G.S. (2012). The Anti-Proliferative Activity of BTG/TOB Proteins Is Mediated via the Caf1a (CNOT7) and Caf1b (CNOT8) Deadenylation Subunits of the Ccr4-Not Complex. *PLoS ONE* 7(12). doi:10.1371/journal.pone.0051331.

79. Maryati M., Airhihen B., & Winkler G. (2015). The enzyme activities of Caf1 and Ccr4 are both required for deadenylation by the human Ccr4-Not nuclease module. *Biochemical Journal* 4922873455(May):Epub ahead of print.
80. Wolf J., Valkov E., Allen M.D., Meineke B., Gordiyenko Y., McLaughlin S.H., Olsen T.M., Robinson C. V, Bycroft M., Stewart M., et al. (2014). Structural basis for Pan3 binding to Pan2 and its function in mRNA recruitment and deadenylation. *The EMBO Journal* 33(14):1514–1526.
81. Uchida N., Hoshino S.-I., & Katada T. (2004). Identification of a human cytoplasmic poly(A) nuclease complex stimulated by poly(A)-binding protein. *The Journal of biological chemistry* 279(2):1383–91.
82. Wolf J., & Passmore L.A. (2014). Deadenylation by Pan2 / Pan3. *Biochem Soc Trans* 42(1):184–187.
83. Tharun S., & Parker R. (2001). Targeting an mRNA for decapping: Displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. *Molecular Cell* 8(5):1075–1083.
84. Fenger-Grøn M., Fillman C., Norrild B., Lykke-Andersen J., Wahle E., Seraphin B., Bertrand E., Filipowicz W., Golan D.E., & Anderson P. (2005). Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. *Molecular cell* 20(6):905–15.
85. Yu J.H., Yang W.-H., Gulick T., Bloch K.D., & Bloch D.B. (2005). Ge-1 is a central component of the mammalian cytoplasmic mRNA processing body. *RNA (New York, NY)* 11(12):1795–802.
86. Morris K. V, & Mattick J.S. (2014). The rise of regulatory RNA. *Nature reviews Genetics* 15(6):423–37.
87. McStay B., & Grummt I. (2008). The Epigenetics of rRNA Genes: From Molecular to Chromosome Biology. *Annual Review of Cell and Developmental Biology* 24(1):131–157.
88. Quinn J.J., & Chang H.Y. (2015). Unique features of long non-coding RNA biogenesis and function. *Nature reviews Genetics* 17(1):47–62.
89. Schlackow M., Nojima T., Gomes T., Dhir A., Carmo-Fonseca M., & Proudfoot N.J. (2017). Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. *Molecular Cell* 65(1):25–38.
90. Claverie J.-M. (2005). Fewer Genes, More Noncoding RNA. *Science* 309(5740). Available at: <http://science.sciencemag.org/content/309/5740/1529.full>

[Accessed June 14, 2017].

91. Mercer T.R., Dinger M.E., & Mattick J.S. (2009). Long non-coding RNAs: insights into functions. *Nature Reviews Genetics* 10(3):155–159.
92. Brown C.J., Ballabio A., Rupert J.L., Lafreniere R.G., Grompe M., Tonlorenzi R., & Willard H.F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 349(6304):38–44.
93. da Rocha S.T., & Heard E. (2017). Novel players in X inactivation: insights into Xist-mediated gene silencing and chromosome conformation. *Nature Publishing Group* 24(3). doi:10.1038/nsmb.3370.
94. Meng X., Li X., Zhang P., Wang J., Zhou Y., & Chen M. (2016). Circular RNA: an emerging key player in RNA world. *Briefings in bioinformatics*:bbw045.
95. Qu S., Yang X., Li X., Wang J., Gao Y., Shang R., Sun W., Dou K., & Li H. (2015). Circular RNA: a new star of noncoding RNAs. *Cancer Letters*:1–8.
96. Ghildiyal M., & Zamore P.D. (2009). Small silencing RNAs: an expanding universe. *Nat Rev Genet* 10(2):94–108.
97. Carthew R.W., & Sontheimer E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136(4):642–655.
98. Griffiths-Jones S., Saini H.K., Van Dongen S., & Enright A.J. (2008). miRBase: Tools for miRNA genomics. *Nucleic Acids Research* 36(SUPPL. 1):154–158.
99. Jima D.D., Zhang J., Jacobs C., Richards K.L., Dunphy C.H., Choi W.W.L., Au W.Y., Srivastava G., Czader M.B., Rizzieri D.A., et al. (2010). Deep sequencing of the small RNA transcriptome of normal and malignant human B cells identifies hundreds of novel miRNAs. *Blood* 116(23). doi:10.1182/blood-2010-05-285403.
100. Berezikov E., Van Tetering G., Verheul M., Van De Belt J., Van Laake L., Vos J., Verloop R., Van De Wetering M., Guryev V., Takada S., et al. (2006). Many novel mammalian miRNA candidates identified by extensive cloning and RAKE analysis. *Genome Research* 16(10):1289–1298.
101. Rodriguez A., Griffiths-Jones S., Ashurst J.L., & Bradley A. (2004). Identification of mammalian miRNA host genes and transcription units. *Genome Research* 14(10 A):1902–1910.

102. Kim Y.-K., & Kim V.N. (2007). Processing of intronic miRNAs. *The EMBO journal* 26(3):775–783.
103. Chang T.-C., Pertea M., Lee S., Salzberg S.L., & Mendell J.T. (2015). Genome-wide annotation of miRNA primary transcript structures reveals novel regulatory mechanisms. *Genome research* 25(9):1401–9.
104. Lee R.C. (1993). The *C. elegans* Heterochronic Gene *lin-4* Encodes Small RNAs with Antisense Complementarity to *lin-14*. *Cell* 75:843–854.
105. Reinhart B.J., Slack F.J., Basson M., Pasquinelli A.E., Bettinger J.C., Rougvie A.E., Horvitz H.R., & Ruvkun G. (2000). The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403(6772):901–906.
106. Ambros V. (2001). miRNAs: Tiny regulators with great potential. *Cell* 107(7):823–826.
107. Lau, Nelson C.; Lim, Lee P.; Weinstein, Earl G.; Bartel D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294(October):858–862.
108. Lee R.C., Ambros V., Erdmann V.A., Lee R.C., Feinbaum R.L., Ambros V., Reinhart B., Wightman B., Ha I., Ruvkun G., et al. (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science (New York, NY)* 294(5543):862–4.
109. Lagos-Quintana M., Rauhut R., Lendeckel W., & Tuschl T. (2001). Identification of Novel Genes Coding for Small Expressed RNAs. *Science* 294(5543):853–858.
110. Smallridge R. (2001). A small fortune. *Nature reviews Molecular cell biology* 2(12):867.
111. Lu M., Zhang Q., Deng M., Miao J., Guo Y., Gao W., & Cui Q. (2008). An analysis of human miRNA and disease associations. *PLoS ONE* 3(10):1–5.
112. Lee Y., Kim M., Han J., Yeom K.-H., Lee S., Baek S.H., & Kim V.N. (2004). MiRNA genes are transcribed by RNA polymerase II. *The EMBO journal* 23(20):4051–4060.
113. Cai X., Hagedorn C.H., & Cullen B.R. (2004). Human miRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA (New York, NY)* 10(12):1957–1966.
114. Gromak N., Dienstbier M., Macias S., Plass M., Eyraes E., Cáceres J.F., & Proudfoot N.J. (2013). *Drosha* regulates gene expression

- independently of RNA cleavage function. *Cell Reports* 5(6):1499–1510.
115. Gregory R.I., Yan K.-P., Amuthan G., Chendrimada T., Doratotaj B., Cooch N., & Shiekhattar R. (2004). The Microprocessor complex mediates the genesis of miRNAs. *Nature* 432(7014):235–240.
 116. Han J., Lee Y., Yeom K.H., Kim Y.K., Jin H., & Kim V.N. (2004). The Drosha-DGCR8 complex in primary miRNA processing. *Genes and Development* 18(24):3016–3027.
 117. Yeom K.-H.H., Lee Y., Han J., Suh M.R., & Kim V.N. (2006). Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. *Nucleic Acids Research* 34(16):4622–4629.
 118. Auyeung V.C., Ulitsky I., McGeary S.E., & Bartel D.P. (2013). Beyond secondary structure: Primary-sequence determinants license Pri-miRNA hairpins for processing. *Cell* 152(4):844–858.
 119. Alarco C.R. (2015). N6-methyladenosine marks primary miRNAs for processing. *Nature* 519:482–487.
 120. Zhang H., Kolb F. a., Jaskiewicz L., Westhof E., & Filipowicz W. (2004). Single processing center models for human Dicer and bacterial RNase III. *Cell* 118(1):57–68.
 121. Zeng Y., & Cullen B.R. (2006). Recognition and cleavage of primary miRNA transcripts. *Methods in molecular biology* (Clifton, NJ) 342(1):49–56.
 122. Nguyen T.A.A., Jo M.H.H., Choi Y.-G., Park J., Kwon S.C.C., Hohng S., Kim V.N.N., & Woo J.-S. (2015). Functional Anatomy of the Human Microprocessor. *Cell* 161(6):1374–1387.
 123. Han J., Lee Y., Yeom K.H., Nam J.W., Heo I., Rhee J.K., Sohn S.Y., Cho Y., Zhang B.T., & Kim V.N. (2006). Molecular Basis for the Recognition of Primary miRNAs by the Drosha-DGCR8 Complex. *Cell* 125(5):887–901.
 124. Han J., Pedersen J.S., Kwon S.C., Belair C.D., Kim Y.-K.K., Yeom K.-H.H., Yang W.-Y.Y., Haussler D., Blelloch R., & Kim V.N. (2009). Posttranscriptional Crossregulation between Drosha and DGCR8. *Cell* 136(1):75–84.
 125. Triboulet R., Chang H.-M., Lapierre R.J., & Gregory R.I. (2009). Post-transcriptional control of DGCR8 expression by the Microprocessor. *RNA* (New York, NY) 15(6):1005–1011.
 126. Herbert K.M., Pimienta G., DeGregorio S.J., Alexandrov A., & Steitz

- J.A. (2013). Phosphorylation of DGCR8 Increases Its Intracellular Stability and Induces a Progrowth miRNA Profile. *Cell Reports* 5(4):1070–1081.
127. Zhu C., Chen C., Huang J., Zhang H., Zhao X., Deng R., Dou J., Jin H., Chen R., Xu M., et al. (2015). SUMOylation at K707 of DGCR8 controls direct function of primary miRNA. *Nucleic acids research* 43(16):gkv741-.
 128. Zhu C., Chen C., Chen R., Deng R., Zhao X., Zhang H., Duo J., Chen Q., Jin H., Wang Y., et al. (2016). K²⁵⁹-SUMOylation of DGCR8 promoted by p14ARF exerts a tumor-suppressive function. *Journal of Molecular Cell Biology* 8(5):456–458.
 129. Herbert K.M., Sarkar S.K., Mills M., Delgado De la Herran H.C., Neuman K.C., & Steitz J.A. (2015). A heterotrimer model of the complete Microprocessor complex revealed by single-molecule subunit counting. *RNA (New York, NY)* 22(2):175–183.
 130. Quick-Cleveland J., Jacob J.P., Weitz S.H., Shoffner G., Senturia R., & Guo F. (2014). The DGCR8 RNA-Binding Heme Domain Recognizes Primary MiRNAs by Clamping the Hairpin. *Cell Reports* 7(6):1994–2005.
 131. Barr I., Weitz S.H.H., Atkin T., Hsu P., Karayiorgou M., Gogos J.A.A., Weiss S., & Guo F. (2015). Cobalt(III) Protoporphyrin Activates the DGCR8 Protein and Can Compensate miRNA Processing Deficiency. *Chemistry & Biology* 22(6):793–802.
 132. Weitz S.H., Gong M., Barr I., Weiss S., & Guo F. (2014). Processing of miRNA primary transcripts requires heme in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* 111(5):1861–6.
 133. Morlando M., Ballarino M., Gromak N., Pagano F., Bozzoni I., & Proudfoot N.J. (2008). Primary miRNA transcripts are processed co-transcriptionally. *Nature structural & molecular biology* 15(9):902–909.
 134. Ballarino M., Pagano F., Girardi E., Morlando M., Cacchiarelli D., Marchioni M., Proudfoot N.J., & Bozzoni I. (2009). Coupled RNA processing and transcription of intergenic primary miRNAs. *Molecular and cellular biology* 29(20):5632–5638.
 135. Szcześniak M.W., & Makałowska I. (2016). lncRNA-RNA Interactions across the Human Transcriptome. *PloS one* 11(3):e0150353.
 136. Dai L., Chen K., Youngren B., Kulina J., Yang A., Guo Z., Li J., Yu P.,

- & Gu S. (2016). Cytoplasmic Drosha activity generated by alternative splicing. *Nucleic acids research*:1–13.
137. Yi R., Qin Y., Macara I.G., & Cullen B.R. (2003). Exportin-5 mediates the nuclear export of pre-miRNAs and short hairpin RNAs. *Genes and Development* 17(24):3011–3016.
 138. Kim Y.-K., Kim B., & Kim V.N. (2016). Re-evaluation of the roles of DROSHA, Export in 5, and DICER in miRNA biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 113(13):E1881-9.
 139. Bernstein E., Caudy a a, Hammond S.M., & Hannon G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409(6818):363–366.
 140. MacRae I.J., Zhou K., & Doudna J. a (2007). Structural determinants of RNA recognition and cleavage by Dicer. *Nature structural & molecular biology* 14(10):934–940.
 141. Preall J.B., & Sontheimer E.J. (2005). RNAi: RISC gets loaded. *Cell* 123(4):543–545.
 142. Wilson R.C., & Doudna J.A. (2013). Molecular Mechanisms of RNA Interference. *Annual Review of Biophysics* 42(1):217–239.
 143. Gregory R.I., Chendrimada T.P., Cooch N., & Shiekhattar R. (2005). Human RISC couples miRNA biogenesis and posttranscriptional gene silencing. *Cell* 123(4):631–640.
 144. Wang H.-W., Noland C., Siridechadilok B., Taylor D.W., Ma E., Felderer K., Doudna J. a, & Nogales E. (2009). Structural insights into RNA processing by the human RISC-loading complex. *Nature structural & molecular biology* 16(11):1148–1153.
 145. Ha M., & Kim V.N. (2014). Regulation of miRNA biogenesis. *Nature reviews Molecular cell biology* 15(8):509–524.
 146. Jagadeeswaran G., Zheng Y., Sumathipala N., Jiang H., Arrese E.L., Soulages J.L., Zhang W., & Sunkar R. (2010). Deep sequencing of small RNA libraries reveals dynamic regulation of conserved and novel miRNAs and miRNA-stars during silkworm development. *BMC Genomics* 11(52):1–18.
 147. Meister G., Landthaler M., Patkaniowska A., Dorsett Y., Teng G., & Tuschl T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Molecular Cell* 15(2):185–197.
 148. Chendrimada T.P., Gregory R.I., Kumaraswamy E., Norman J., Cooch N., Nishikura K., & Shiekhattar R. (2005). TRBP recruits the

- Dicer complex to Ago2 for miRNA processing and gene silencing. *Nature* 436(7051):740–744.
149. Lee H.Y., Zhou K., Smith A.M., Noland C.L., & Doudna J. a. (2013). Differential roles of human Dicer-binding proteins TRBP and PACT in small RNA processing. *Nucleic Acids Research* 41(13):6568–6576.
 150. Khvorova A., Reynolds A., & Jayasena S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115(2):209–216.
 151. Winter J., & Diederichs S. (2013). Argonaute-3 activates the let-7a passenger strand miRNA. *RNA biology* 10(10):1631–43.
 152. Salomon W.E., Jolly S.M., Moore M.J., Zamore P.D., & Serebrov V. (2015). Single-Molecule Imaging Reveals that Argonaute Reshapes the Binding Properties of Its Nucleic Acid Guides. *Cell* 162(1):84–95.
 153. Schirle N.T., Sheu-Gruttadauria J., & MacRae I.J. (2014). Structural basis for miRNA targeting. *Science* 346(6209):608–613.
 154. Chandradoss S.D., Schirle N.T., Szczepaniak M., MacRae I.J., & Joo C. (2015). A Dynamic Search Process Underlies MiRNA Targeting. *Cell* 162(1):96–107.
 155. Jiang H., Sheong F.K., Zhu L., Gao X., Bernauer J., & Huang X. (2015). Markov State Models Reveal a Two-Step Mechanism of miRNA Loading into the Human Argonaute Protein: Selective Binding followed by Structural Re-arrangement. *PLoS computational biology* 11(7):e1004404.
 156. Yao B., La L.B., Chen Y.-C., Chang L.-J., & Chan E.K.L. (2012). Defining a new role of GW182 in maintaining miRNA stability. *EMBO reports* 13(12):1102–8.
 157. Ricci E.P., Limousin T., Soto-Rifo R., Allison R., Pöyry T., Decimo D., Jackson R.J., & Ohlmann T. (2011). Activation of a miRNA response in trans reveals a new role for poly(A) in translational repression. *Nucleic Acids Research* 39(12):5215–5231.
 158. Fabian M.R., Mathonnet G., Sundermeier T., Mathys H., Zipprich J.T., Svitkin Y. V, Rivas F., Jinek M., Wohlschlegel J., Doudna J.A., et al. (2009). Mammalian miRNA RISC Recruits CAF1 and PABP to Affect PABP-Dependent Deadenylation. *Molecular Cell* 35(6):868–880.
 159. Chen C.-Y.A., Zheng D., Xia Z., & Shyu A.-B. (2010). Ago-TNRC6 complex triggers miRNA-mediated mRNA decay by promoting biphasic deadenylation followed by decapping. *Nature structural & molecular biology* 16(11):1160–1166.

160. Chekulaeva M., Mathys H., Zipprich J.T., Attig J., Colic M., Parker R., & Filipowicz W. (2011). miRNA repression involves GW182-mediated recruitment of CCR4–NOT through conserved W-containing motifs. *Nature Structural & Molecular Biology* 18(11):1218–1226.
161. Piao X., Zhang X., Wu L., & Belasco J.G. (2010). CCR4–NOT deadenylates mRNA associated with RNA-induced silencing complexes in human cells. *Molecular and cellular biology* 30(6):1486–1494.
162. Takimoto K., Wakiyama M., & Yokoyama S. (2009). Mammalian GW182 contains multiple Argonaute-binding sites and functions in miRNA-mediated translational repression. *RNA (New York, NY)* 15(6):1078–1089.
163. Fabian M.R., Cieplak M.K., Frank F., Morita M., Green J., Srikumar T., Nagar B., Yamamoto T., Raught B., Duchaine T.F., et al. (2012). miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4–NOT. *Nature Structural & Molecular Biology* 19(3):1211–1217.
164. Nishihara T., Zekri L., Braun J.E., & Izaurralde E. (2013). MiRISC recruits decapping factors to miRNA targets to enhance their degradation. *Nucleic Acids Research* 41(18):8692–8705.
165. Wilczynska A., & Bushell M. (2015). The complexity of miRNA-mediated repression. *Cell Death and Differentiation* 22(1):22–33.
166. Fabian M.R., & Sonenberg N. (2012). The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. *Nature Structural & Molecular Biology* 19(6):586–593.
167. Meyer K.D., Saletore Y., Zumbo P., Elemento O., Mason C.E., & Jaffrey S.R. (2012). Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3' UTRs and Near Stop Codons. *Cell* 149(7):1635–1646.
168. Laubinger S., Sachsenberg T., Zeller G., Busch W., Lohmann J.U., Räscher G., & Weigel D. (2008). Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and miRNA processing in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* 105(25):8795–8800.
169. Gruber J.J., Zatechka D.S., Sabin L.R., Yong J., Lum J.J., Kong M., Zong W.X., Zhang Z., Lau C.K., Rawlings J., et al. (2009). Ars2 Links the Nuclear Cap-Binding Complex to RNA Interference and Cell Proliferation. *Cell* 138(2):328–339.

170. Sabin L.R., Zhou R., Gruber J.J., Lukinova N., Bambina S., Berman A., Lau C.K., Thompson C.B., & Cherry S. (2009). Ars2 Regulates Both miRNA- and siRNA- Dependent Silencing and Suppresses RNA Virus Infection in Drosophila. *Cell* 138(2):340–351.
171. Morlando M., Dini Modigliani S., Torrelli G., Rosa A., Di Carlo V., Caffarelli E., & Bozzoni I. (2012). FUS stimulates miRNA biogenesis by facilitating co-transcriptional Drosha recruitment. *The EMBO Journal* 31(24):4502–4510.
172. Tang X., Li M., Tucker L., & Ramratnam B. (2011). Glycogen synthase kinase 3 beta (GSK3 β) phosphorylates the RNAase III enzyme Drosha at S300 and S302. *PLoS ONE* 6(6):1–6.
173. Tang X., Zhang Y., Tucker L., & Ramratnam B. (2010). Phosphorylation of the RNase III enzyme drosha at Serine300 or Serine302 is required for its nuclear localization. *Nucleic Acids Research* 38(19):6610–6619.
174. Fletcher C.E., Godfrey J.D., Shibakawa A., Bushell M., & Bevan C.L. (2016). A novel role for GSK3 β as a modulator of Drosha microprocessor activity and MiRNA biogenesis. *Nucleic Acids Research* 42(5):gkw938.
175. Moy R.H., Cole B.S., Yasunaga A., Gold B., Shankarling G., Varble A., Molleston J.M., TenOever B.R., Lynch K.W., & Cherry S. (2014). Stem loop recognition by DDX17 facilitates miRNA processing and antiviral defense. *Cell* 158(4):764–777.
176. Fukuda T., Yamagata K., Fujiyama S., Matsumoto T., Koshida I., Yoshimura K., Mihara M., Naitou M., Endoh H., Nakamura T., et al. (2007). DEAD-box RNA helicase subunits of the Drosha complex are required for processing of rRNA and a subset of miRNAs. *Nature Cell Biology* 9(5):604–611.
177. Dardenne E., PolayEspinoza M., Fattet L., Germann S., Lambert M.P., Neil H., Zonta E., Mortada H., Gratadou L., Deygas M., et al. (2014). RNA Helicases DDX5 and DDX17 Dynamically Orchestrate Transcription, miRNA, and Splicing Programs in Cell Differentiation. *Cell Reports* 7(6):1900–1913.
178. Connerty P., Bajan S., Remenyi J., Fuller-Pace F. V., & Hutvagner G. (2016). The miRNA biogenesis factors, p72/DDX17 and KHSRP regulate the protein level of Ago2 in human cells. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1859(10):1299–1305.

179. Garibaldi F., Falcone E., Trisciuoglio D., Colombo T., Lisek K., Walerych D., Del Sal G., Paci P., Bossi G., Piaggio G., et al. (2016). Mutant p53 inhibits miRNA biogenesis by interfering with the microprocessor complex. *Oncogene* 35(January):3760–3770.
180. Mori M., Triboulet R., Mohseni M., Schlegelmilch K., Shrestha K., Camargo F.D., & Gregory R.I. (2014). Hippo signaling regulates microprocessor and links cell-density-dependent mirna biogenesis to cancer. *Cell* 156(5):893–906.
181. Bhattacharyya S.N., Habermacher R., Martine U., Closs E.I., & Filipowicz W. (2006). Relief of miRNA-Mediated Translational Repression in Human Cells Subjected to Stress. *Cell* 125(6):1111–1124.
182. Libri V., Miesen P., Van Rij R.P., & Buck A.H. (2013). Regulation of miRNA biogenesis and turnover by animals and their viruses. *Cellular and Molecular Life Sciences* 70(19):3525–3544.
183. Rüegger S., & Großhans H. (2012). MiRNA turnover: when, how, and why. *Trends in biochemical sciences* 37(10):436–46.
184. Yu B., Yang Z., Li J., Minakhina S., Yang M., Padgett R.W., Steward R., & Chen X. (2005). Methylation as a Crucial Step in Plant miRNA Biogenesis. *Science* 307(5711):932–935.
185. Yang Z., Ebright Y.W., Yu B., & Chen X. (2006). HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Research* 34(2):667–675.
186. Li J., Yang Z., Liu J., & Chen X. (2005). Methylation Protects miRNAs and siRNAs from a 3'-End Uridylation Activity in Arabidopsis. *Curr Biol* 15(16):1501–1507.
187. Zhao Y., Yu Y., Zhai J., Ramachandran V., & Theresa T. (2012). HESO1, a nucleotidyl transferase in Arabidopsis, uridylates unmethylated miRNAs and siRNAs to trigger their degradation. *Curr Biol* 22(8):689–694.
188. Katoh T., Sakaguchi Y., Miyauchi K., Suzuki T., Suzuki T., Kashiwabara S.I., & Baba T. (2009). Selective stabilization of mammalian miRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes and Development* 23(4):433–438.
189. Rissland O.S., Hong S.-J., & Bartel D.P. (2011). miR-16 Family in Response to Cell Cycle Changes. *Mol Cell* 43(6):993–1004.
190. Zhang Z., Zou J., Wang G.K., Zhang J.T., Huang S., Qin Y.W., & Jing

- Q. (2011). Uracils at nucleotide position 9-11 are required for the rapid turnover of miR-29 family. *Nucleic Acids Research* 39(10):4387–4395.
191. Ameres S.L., Horwich M.D., Hung J., Xu J., Ghildiyal M., Weng Z., & Zamore P.D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. *Science* 328(5985):1534–1539.
 192. Saini H.K., Griffiths-Jones S., & Enright A.J. (2007). Genomic analysis of human miRNA transcripts. *Proceedings of the National Academy of Sciences of the United States of America* 104(45):17719–17724.
 193. Kataoka N., Fujita M., & Ohno M. (2009). Functional association of the Microprocessor complex with the spliceosome. *Molecular and cellular biology* 29(12):3243–3254.
 194. Kim Y., & Kim V.N. (2012). MiRNA Factory: RISC Assembly from Precursor MiRNAs. *Molecular Cell* 46(4):384–386.
 195. Agranat-Tamir L., Shomron N., Sperling J., & Sperling R. (2014). Interplay between pre-mRNA splicing and miRNA biogenesis within the supraspliceosome. *Nucleic Acids Research* 42(7):4640–4651.
 196. Mondol V., Ahn B.C., & Pasquinelli A.M.Y.E. (2015). Splicing remodels the let-7 primary miRNA to facilitate Drosha processing in *Caenorhabditis elegans*. *RNA (New York, NY)* 21:1–8.
 197. Sundaram G.M., Common J.E. a, Gopal F.E., Srikanta S., Lakshman K., Lunny D.P., Lim T.C., Tanavde V., Lane E.B., & Sampath P. (2013). “See-saw” expression of miRNA-198 and FSTL1 from a single transcript in wound healing. *Nature* 495(7439):103–6.
 198. Mattioli C., Pianigiani G., & Pagani F. (2013). A competitive regulatory mechanism discriminates between juxtaposed splice sites and pri-miRNA structures. *Nucleic Acids Research* 41(18):8680–8691.
 199. Ruby J.G., Jan C.H., & Bartel D.P. (2007). Intronic miRNA precursors that bypass Drosha processing. *Nature* 448(7149):83–86.
 200. Xie M., Li M., Vilborg A., Lee N., Shu M. Di, Yartseva V., Šestan N., & Steitz J. a. (2013). Mammalian 5'-capped miRNA precursors that generate a single miRNA. *Cell* 155(7):1568–1580.
 201. Cheloufi S., Dos Santos C.O., Chong M.M.W., & Hannon G.J. (2010). A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* 465(7298):584–589.
 202. Knuckles P., Carl S.H., Musheev M., Niehrs C., Wenger A., & Bühler

- M. (2017). RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. *Nature Publishing Group*. doi:10.1038/nsmb.3419.
203. Kim B., Jeong K., & Kim V.N. (2017). Genome-wide Mapping of DROSHA Cleavage Sites on Primary MiRNAs and Noncanonical Substrates. *Molecular Cell* 66(2):258–269.e5.
 204. Cirera-Salinas D., Yu J., Bodak M., Ngondo R.P., Herbert K.M., & Ciaudo C. (2017). Noncanonical function of DGCR8 controls mESC exit from pluripotency. *The Journal of Cell Biology* 216(2):355–366.
 205. Heras S.R., Macias S., Plass M., Fernandez N., Cano D., Eyraes E., Garcia-Perez J.L., & Cáceres J.F. (2013). The Microprocessor controls the activity of mammalian retrotransposons. *Nature structural & molecular biology* 20(10):1173–81.
 206. Dhir A., Dhir S., Proudfoot N.J., & Jopling C.L. (2015). Microprocessor mediates transcriptional termination in genes encoding long noncoding miRNAs. *Nature Structural & Molecular Biology* 2015(November 2014). doi:10.1038/nsmb.2982.
 207. Lagos-Quintana M., Rauhut R., Yalcin A., Meyer J., Lendeckel W., & Tuschl T. (2002). Identification of tissue-specific MiRNAs from mouse. *Current Biology* 12(9):735–739.
 208. Coulouarn C., Factor V.M., Andersen J.B., Durkin M.E., & Thorgeirsson S.S. (2009). Loss of miR-122 expression in liver cancer correlates with suppression of the hepatic phenotype and gain of metastatic properties. *Oncogene* 28(40):3526–3536.
 209. Bandiera S., Pfeffer S., Baumert T.F., & Zeisel M.B. (2015). miR-122 – A key factor and therapeutic target in liver disease. *Journal of Hepatology* 62(2):448–457.
 210. Jopling C.L., Yi M., Lancaster A.M., Lemon S.M., & Sarnow P. (2005). Modulation of hepatitis C virus RNA abundance by a liver-specific MiRNA. *Science (New York, NY)* 309(5740):1577–1581.
 211. Lewis A.P., & Jopling C.L. (2010). Regulation and biological function of the liver-specific miR-122. *Biochemical Society transactions* 38(6):1553–1557.
 212. Cox E.M., Sagan S.M., Mortimer S. a W., Doudna J. a, & Sarnow P. (2013). Enhancement of hepatitis C viral RNA abundance by precursor miR-122 molecules. *RNA (New York, NY)* 19(12):1825–32.
 213. Jopling C.L., Schütz S., & Sarnow P. (2008). Position-Dependent Function for a Tandem MiRNA miR-122-Binding Site Located in the

- Hepatitis C Virus RNA Genome. *Cell Host and Microbe* 4(1):77–85.
214. Lanford R.E., Hildebrandt-eriksen E.S., Petri A., Persson R., Lindow M., Munk M.E., Kauppinen S., & Ørum H. (2010). Therapeutic silencing of miRNA-122 in primates with chronic hepatitis C virus infection. *Science* 327(5962):198–201.
 215. Jopling C.L. (2010). Targeting miRNA-122 to treat hepatitis C virus infection. *Viruses* 2(7):1382–1393.
 216. Van Der Ree M.H., Van Der Meer A.J., Van Nuenen A.C., De Bruijne J., Ottosen S., Janssen H.L., Kootstra N.A., & Reesink H.W. (2016). Miravirsin dosing in chronic hepatitis C patients results in decreased miRNA-122 levels without affecting other miRNAs in plasma. *Alimentary Pharmacology and Therapeutics* 43(1):102–113.
 217. Tanzer A., & Stadler P.F. (2004). Molecular Evolution of a MiRNA Cluster. *Journal of Molecular Biology* 339(2):327–335.
 218. Chaulk S.G., Thede G.L., Kent O.A., Xu Z., Gesner E., Veldhoen R.A., Khanna S.K., Goping I.S., MacMillan A.M., Mendell J.T., et al. (2011). Role of pri-miRNA tertiary structure in miR-17~92 miRNA biogenesis. *RNA Biology*. doi:10.4161/rna.8.6.17410.
 219. Du P., Wang L., Sliz P., & Gregory R.I. (2015). A Biogenesis Step Upstream of Microprocessor Controls miR-17~92 Expression. *Cell* 162(4):885–899.
 220. Mogilyansky E., & Rigoutsos I. (2013). The miR-17/92 cluster: a comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. *Cell Death and Differentiation* 20(12):1603–1614.
 221. Ventura A., Young A.G., Winslow M.M., Lintault L., Meissner A., Erkeland S.J., Newman J., Bronson R.T., Crowley D., Stone J.R., et al. (2008). Targeted Deletion Reveals Essential and Overlapping Functions of the miR-17~92 Family of miRNA Clusters. *Cell* 132(5):875–886.
 222. Cho S., Wu C.-J., Yasuda T., Cruz L.O., Khan A.A., Lin L.-L., Nguyen D.T., Miller M., Lee H.-M., Kuo M.-L., et al. (2016). miR-23~27~24 clusters control effector T cell differentiation and function. *The Journal of experimental medicine* 213(2):235–49.
 223. Chhabra R., Dubey R., & Saini N. (2010). Cooperative and individualistic functions of the miRNAs in the miR-23a~27a~24-2 cluster and its implication in human diseases. *Molecular cancer* 9(1):232.

224. Feng Y., Chan H., Wang N., Zhu M., Cheung F., & Hong M. (2013). MIR23A (miRNA 23a). *Atlas of Genetics and Cytogenetics in Oncology and Haematology* 17(12):825–827.
225. Ramalingam P., Palanichamy J.K., Singh A., Das P., Bhagat M., Kassab M.A., Sinha S., & Chattopadhyay P. (2014). Biogenesis of intronic miRNAs located in clusters by independent transcription and alternative splicing. *RNA (New York, NY)* 20(1):76–87.
226. Chang T.C., Yu D., Lee Y.S., Wentzel E.A., Arking D.E., West K.M., Dang C. V, Thomas-Tikhonenko A., & Mendell J.T. (2008). Widespread miRNA repression by Myc contributes to tumorigenesis. *Nat Genet* 40(1):43–50.
227. Kriegel A.J., Liu Y., Fang Y., Ding X., & Liang M. (2012). The miR-29 family: genomics, cell biology, and relevance to renal and cardiovascular injury. *Physiological genomics* 44(4):237–44.
228. Roberts A.P.E., Doidge R., Tarr A.W., & Jopling C.L. (2014). The P body protein LSm1 contributes to stimulation of hepatitis C virus translation, but not replication, by miRNA-122. *Nucleic Acids Research* 42(2):1257–1269.
229. Dye M.J., & Proudfoot N.J. (1999). Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. *Molecular Cell* 3(3):371–378.
230. Harwig A., Jongejan A., van Kampen A.H.C., Berkhout B., & Das A.T. (2016). Tat-dependent production of an HIV-1 TAR-encoded miRNA-like small RNA. *Nucleic acids research:gkw167*.
231. Pall G.S., & Hamilton A.J. (2008). Improved northern blot method for enhanced detection of small RNA. *Nature protocols* 3(6):1077–1084.
232. Matera A.G., & Wang Z. (2014). A day in the life of the spliceosome. *Nature reviews Molecular cell biology* 15(2):108–21.
233. Adams S.E., Johnson I.D., Braddock M., Kingsman A.J., Kingsman S.M., & Edwards R.M. (1988). Synthesis of a gene for the HIV transactivator protein TAT by a novel single stranded approach involving in vivo gap repair. *Nucleic Acids Research* 16(10):4287–4298.
234. Janas M.M., Khaled M., Schubert S., Bernstein J.G., Golan D., Veguilla R. a., Fisher D.E., Shomron N., Levy C., & Novina C.D. (2011). Feed-forward microprocessing and splicing activities at a miRNA-containing intron. *PLoS Genetics* 7(10):4–13.

235. Gao Y., Vogt A., Forsyth C.J., & Koide K. (2013). Comparison of Splicing Factor 3b Inhibitors in Human Cells. *ChemBioChem* 14(1):49–52.
236. Kotake Y., Sagane K., Owa T., Mimori-Kiyosue Y., Shimizu H., Uesugi M., Ishihama Y., Iwata M., & Mizui Y. (2007). Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nature chemical biology* 3(9):570–575.
237. Kaida D., Motoyoshi H., Tashiro E., Nojima T., Hagiwara M., Ishigami K., Watanabe H., Kitahara T., Yoshida T., Nakajima H., et al. (2007). Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nature chemical biology* 3(9):576–583.
238. Gatfield D., Martelot G. Le, Vejnar C.E., Gerlach D., Schaad O., Oresic M., Esau C.C., Fleury-olela F., & Ruskeepa A. (2009). Integration of miRNA miR-122 in hepatic circadian gene expression. *Genes and Development* 23:1313–1326.
239. Spiluttini B., Gu B., Belagal P., Smirnova A.S., Nguyen V.T., Hébert C., Schmidt U., Bertrand E., Darzacq X., & Bensaude O. (2010). Splicing-independent recruitment of U1 snRNP to a transcription unit in living cells. *Journal of cell science* 123(Pt 12):2085–2093.
240. Kaida D., Berg M.G., Younis I., Kasim M., Singh L.N., Wan L., & Dreyfuss G. (2010). U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* 468. doi:10.1038/nature09479.
241. Lacadie S.A., & Rosbash M. (2005). Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5'ss base pairing in yeast. *Molecular Cell* 19(1):65–75.
242. Conrad T., Marsico A., Gehre M., Ørom U.A., Ørom U.A., & Ørom U.A. (2014). Microprocessor Activity Controls Differential miRNA Biogenesis In Vivo. *Cell Reports* 9(2):542–554.
243. Dechat T., Adam S.A., Taimen P., Shimi T., & Goldman R.D. (2010). Nuclear lamins. *Cold Spring Harbor perspectives in biology* 2(11):a000547.
244. Luna J.M.M., Scheel T.K.H.K.H., Danino T., Shaw K.S.S., Mele A., Fak J.J.J., Nishiuchi E., Takacs C.N.N., Catanese M.T.T., de Jong Y.P., et al. (2015). Hepatitis C virus RNA functionally sequesters miR-122. *Cell* 160(6):1099–1110.
245. Dye M.J., Gromak N., & Proudfoot N.J. (2006). Exon tethering in transcription by RNA polymerase II. *Molecular Cell* 21(6):849–859.

246. Gatfield D., Martelot G. Le, Vejnar C.E., Gerlach D., Schaad O., Oresic M., Esau C.C., Fleury-olela F., Ruskeepa A., Le Martelot G., et al. (2005). Integration of miRNA miR-122 in hepatic circadian gene expression. *Science (New York, NY)* 309:1577–1581.
247. Rädle B., Rutkowski A.J., Ruzsics Z., Friedel C.C., Koszinowski U.H., & Dölken L. (2013). Metabolic labeling of newly transcribed RNA for high resolution gene expression profiling of RNA synthesis, processing and decay in cell culture. *Journal of visualized experiments : JoVE* (78):1–11.