

**Use of single molecule methods to reveal the
mechanisms of splice site selection**

Is submitted for the degree of
Doctor of Philosophy
at the University of Leicester

September 2011

Mark James Hodson BSc (Hons)

Abstract: Use of single molecule methods to reveal the mechanisms of splice site selection.– Mark Hodson.

Understanding the molecular mechanisms that allow splice site selection is of fundamental importance to the study of alternative splicing. The study presented here has focused on the interactions of two spliceosomal trans-acting factors with pre-mRNA. The U1 snRNP is very important in 5' splice site selection. However, how binding of U1 snRNP to pre-mRNAs containing multiple 5' splice sites leads to 5' splice site selection is unknown. To investigate the role of U1 snRNP in 5' splice site selection, single molecule microscopy was used to determine the number of U1 snRNPs bound to a single molecule of pre-mRNA. In cases where multiple strong 5' splice sites are competing the furthest downstream 5'ss is used. In these cases, as many U1 snRNPs were observed bound to the pre-mRNA at an early stage of the reaction as there were 5' splice sites, demonstrating that multiple sites were occupied. However, later pre-spliceosomal complexes contained only one U1 snRNP, indicating that the surplus U1 snRNP had been removed. A novel model of 5' splice site selection is presented, in which U1snRNP stimulates the binding of SR proteins to the pre-mRNA and thus increases exon rigidity, enforcing use of the downstream 5' splice site.

The U2 Auxiliary Factor (U2AF), binds to the polypyrimidine tract and 3' splice site of an intron and is a heterodimer of the U2AF65 and U2AF35 subunits. The length of the polypyrimidine tract has been shown to influence the efficiency with which a 3' splice site is spliced. To investigate the role of polypyrimidine tract length on U2AF binding, experiments have been conducted in which the binding of U2AF65 to pre-mRNAs has been analysed using single molecule microscopy. Through the analysis of constructs derived from the Globin C pre-mRNA, the mechanisms of U2AF65 association with pre-mRNA have been shown to comprise; (i) a non-specific association that occurs with all RNA tested in the absence of ATP, (ii) a U1 snRNP-dependent association with pre-mRNAs in a polypyrimidine tract-independent manner, and (iii) an association with pre-mRNA in a polypyrimidine tract-dependent manner.

Acknowledgments

I would first like to thank my supervisor, Professor I. C. Eperon for allowing me to study for a PhD within his lab and his support and encouragement throughout. I would like to thank all members of the lab, in particular Dr L.P. Eperon, Dr N. Owen, Miss L. Smith and Mrs R. Dumas for friendly company and scientific discussion. For help and guidance regarding use of the Total-Internal-Reflection-Fluorescence microscope, I would like to thank Dr D. Cherny and Professor C. Bagshaw. For giving me strength and support outside of the laboratory, both during my studies and whilst writing this thesis I thank my wife, mother, father and sister. I also especially thank both Dr L.P. Eperon and Professor I.C. Eperon for reading this thesis.

Contents

	<u>Page</u>
<u>Chapter 1. Introduction</u>	2
<u>1.1 Spliceosome assembly</u>	4
<u>1.2 Splice site selection</u>	14
<u>1.3 Selection of 5' splice sites</u>	15
<u>1.4 Mechanism of 5' splice site selection</u>	21
<u>1.5 Selection of 3' splice sites</u>	26
<u>1.6 Exon definition</u>	34
<u>1.7 The commitment to splice</u>	36
<u>1.8 Total internal reflection fluorescence (TIRF) microscopy</u>	36
<u>Chapter 2. Methods</u>	42
<u>2.1 Procedures related to splicing</u>	45
<u>2.2 Protein analysis techniques</u>	53
<u>2.3 Methods related to total internal reflection fluorescence</u> <u>microscopy</u>	60
<u>2.4 Other techniques</u>	67
<u>2.5 Buffers</u>	70
<u>Chapter 3. Investigations into aspects of splicing: splicing optimization,</u> <u>complex analysis, studies of U1 snRNP and DNase</u>	72
<u>3.1 Optimization of splicing</u>	74
<u>3.2 Characterization of spliceosomal complexes by native</u> <u>agarose gel electrophoresis</u>	78
<u>3.3 A study of the complexes formed by fluorescent fusion</u> <u>proteins in NE</u>	92
<u>3.4 An unusual effect of U1 snRNA knockdown by</u> <u>DNA-mediated RNase H cleavage</u>	98
<u>3.5 The effect of DNase on splicing</u>	107
<u>Chapter 4. Investigations into the number of U1 snRNPs bound to pre-</u> <u>mRNAs containing multiple 5' SS</u>	111
<u>4.1 Assays to determine whether fluorescent fusion proteins</u> <u>are functional</u>	112
<u>4.2 Construction, characterization and labelling of the</u> <u>pre-mRNAs</u>	116
<u>4.3 Preparation of samples for TIRF microscopy</u>	124
<u>4.4 One U1 snRNP associates with pre-mRNAs containing a</u> <u>single 5' SS under E complex conditions</u>	126

4.5 Two U1 snRNPs probably associate with pre-mRNAs containing two 5' SS under E complex conditions	136
4.6 Loss of one U1 snRNP from pre-mRNAs containing two 5' SS by A complex	139
4.7 Loss of U1 snRNP is dependent on the presence of a 3' SS downstream and is inhibited by the presence of a 3' SS between the two 5' SS	146
Chapter 5. U2AF and its interactions with pre-mRNA	154
5.1 Preliminary detection of co-localization between mEGFP-U2AF65 and a globin pre-mRNA	157
5.2 Functional characterization of U2AF65 and U2AF35 fluorescent fusion proteins	162
5.3 Co-localization between mEGFP-U2AF65 and a 'negative control' pre-mRNA	167
5.4 Analysis of the 24-36 U/C pY tract of globin C	172
5.5 An extensive analysis of co-localization efficiencies between mEGFP-U2AF65 and globin C derivatives	180
5.6 Functional characterization of other fluorescent fusion proteins	186
Chapter 6. Discussion	191
6.1 The effect of DNase on globin C pre-mRNA splicing	192
6.2 The role of U1 snRNP in 5' SS selection	194
6.3 Interpretation of data collected from TIRF microscope	209
6.4 Possible explanations for the intensity profiles of mEGFP-U2AF65	212
6.5 Mechanisms of association of mEGFP-U2AF65 with globin C pre-mRNA	213
6.6 Concluding remarks	217
Appendix 1	219
Abbreviations	220
References	222

List of Figures

Chapter 1. Introduction

<u>1.1 Pre-mRNA splicing</u>	<u>3</u>
<u>1.2 The U1 snRNP</u>	<u>7</u>
<u>1.3 Mechanisms of 5' splice site selection based upon affinity</u>	<u>22</u>
<u>1.4 The independent occupation model</u>	<u>24</u>
<u>1.5 Prism based TIRF microscopy</u>	<u>38</u>

Chapter 2. Methods

<u>2.1 The arrangement of the slide for single molecule experiments</u>	<u>62</u>
--	------------------

Chapter 3. Investigations into aspects of splicing: splicing optimization, complex analysis, studies of U1 snRNP and DNase

<u>3.1 Optimization of splicing reactions</u>	<u>75</u>
<u>3.2 ATP depletion mediated by Hexokinase and glucose</u>	<u>80</u>
<u>3.3 Assignment of possible E complex using snRNA degradation</u>	<u>82</u>
<u>3.4 E complex assembly and the effect of phosphatase inhibition</u>	<u>85</u>
<u>3.5 Characterization of A complex</u>	<u>87</u>
<u>3.6 Characterization of the spliceosome using RNase H – mediated snRNA degradation</u>	<u>89</u>
<u>3.7 Confirmation of the presence of B^{ACT} and C complexes</u>	<u>91</u>
<u>3.8 The effect of phosphatase inhibition on heparin-resistant complexes</u>	<u>93</u>
<u>3.9 Formation of complexes by mEGFP-U1-A in NE</u>	<u>95</u>
<u>3.10 Detection of snRNA using Northern blotting</u>	<u>97</u>
<u>3.11 Detection of unusual products when U1snRNA was degraded using RNase H</u>	<u>100</u>
<u>3.12 A comparison between the effects of knockdown and sequestration of U1 snRNP on splicing efficiency</u>	<u>101</u>
<u>3.13 Characterization of the U1 snRNA knockdown-dependent products</u>	<u>104</u>
<u>3.14 The products formed in a U1 knockdown reaction are not related to splicing</u>	<u>106</u>
<u>3.15 The DNase effect</u>	<u>109</u>

Chapter 4. Investigations into the number of U1 snRNPs bound to pre-mRNAs containing multiple 5' SS

<u>4.1 The mEGFP-U1-A fluorescent fusion protein is functional</u>	115
<u>4.2 Representation of pre-mRNAs</u>	117
<u>4.3 Labelling of the RNA</u>	122
<u>4.4 Characterization of NE made from mEGFP-U1-A transfected HeLa cells</u>	127
<u>4.5 Examples of co-localization data</u>	129
<u>4.6 Representative single step intensity profiles</u>	131
<u>4.7 Representative two step intensity profiles</u>	132
<u>4.8 Representative unassignable intensity profiles</u>	134
<u>4.9 Altering the 5' SS selection on the G176G pre-mRNA by addition of the α-U1 oligonucleotide</u>	140
<u>4.10 Spliceosomal assembly on the C24C pre-mRNA</u>	144
<u>4.11 Spliceosome assembly on CEC pre-mRNA</u>	150
<u>4.12 Characterization of the spliceosome assembly on CEC pre-mRNA</u>	151

Chapter 5. U2AF and its interactions with pre-mRNA

<u>5.1 Example of co-localization data for C174C and mEGFP-U2AF65</u>	158
<u>5.2 Transfection of U2AF65 and U2AF35 constructs into HEK 293T cells</u>	161
<u>5.3 Splicing and complex formation in HEK 293T cells</u>	163
<u>5.4 Characterization of NE</u>	165
<u>5.5 Functional assays for mEGFP-U2AF65 and mEGFP-U2AF35</u>	168
<u>5.6 Example co-localization data for mEGFP-U2AF65</u>	171
<u>5.7 Deletion of the 24-36 U/C region</u>	173
<u>5.8 Characterization of globin C Δ24-36U/C</u>	177
<u>5.9 Co-localization between globin C ΔpY,Δ24-36U/C and mEGFP-U2AF65</u>	179
<u>5.10 Characterization of the NE used for TIRF microscopy experiments</u>	181
<u>5.11 Functional assays performed on other spliceosomal factors</u>	188

Chapter 6. Discussion

<u>6.1 A model for the selection of 5' SS</u>	210
--	------------

<u>6.2 Possible mechanisms of U2AF65 binding</u>	<u>216</u>
<u>6.3n Possible role of U1 snRNP in the association of U2AF65</u>	
<u>with pre-mRNA</u>	<u>216</u>

List of Tables

Chapter 1. Introduction

<u>1.1 Protein and RNA composition of the human snRNPs</u>	<u>5</u>
---	-----------------

Chapter 2. Methods

<u>2.1 Sequences of oligonucleotides</u>	<u>46</u>
<u>2.2 Band sizes of pBR322 digests</u>	<u>51</u>
<u>2.3 Primary and secondary antibodies</u>	<u>55</u>
<u>2.4 Laser and emission filter requirements for fluorophores</u>	<u>63</u>
<u>2.5 Dichroic mirrors used to separate emission wavelengths</u>	<u>63</u>

Chapter 3. Investigations into aspects of splicing: splicing optimization, complex analysis, studies of U1 snRNP and DNase

<u>3.1 Spliceosomal complexes and their requirements</u>	<u>78</u>
<u>3.2 Sequences of oligonucleotides used for RNase H degradation</u>	<u>81</u>
<u>3.3 Sequences of oligonucleotides used for site directed mutagenesis</u>	<u>83</u>

Chapter 4. Investigations into the number of U1 snRNPs bound to pre-mRNAs containing multiple 5' SS

<u>4.1 Sequences of oligonucleotides</u>	<u>119</u>
<u>4.2 Sequences of fluorescent oligonucleotides</u>	<u>125</u>
<u>4.3 Quantification of the number of mEGFP-U1-A in spots co-localized with the Anti-U1-oligonucleotide or non co-localized</u>	<u>135</u>
<u>4.4 Approximate efficiency of co-localization between mEGFP-U1-A and Cy5 labelled pre-mRNA</u>	<u>136</u>
<u>4.5 Proportion of mEGFP-U1-A spots which bleached in one and two steps when under E complex conditions</u>	<u>138</u>
<u>4.6 Proportion of mEGFP-U1-A spots which bleached in one and two steps when under A complex conditions</u>	<u>142</u>
<u>4.7 Co-localization between Cy5 and Cy3- labelled oligonucleotide attached to Ad1CC pre-mRNA</u>	<u>142</u>
<u>4.8 Proportion of mEGFP-U1-A spots which bleached in one and two steps when co-localized with Ad1CC pre and post ATP addition</u>	<u>146</u>
<u>4.9 Proportion of mEGFP-U1-A spots which bleached in one</u>	

<u>and two steps when co-localized with CEC</u>	152
---	-----

Chapter 5. U2AF and its interactions with pre-mRNA

<u>5.1 Oligonucleotides used for the construction of pre-mRNA templates</u>	170
<u>5.2 Efficiency of co-localization between globin CΔpY,Δ24-36U/C and mEGFP-U2AF65</u>	180
<u>5.3 Efficiency of co-localization between mEGFP-U2AF65 and globin C or globin M</u>	183
<u>5.4 Efficiency of co-localization between mEGFP-U2AF65 and globin derivatives lacking the 3' SS region</u>	185
<u>5.5 Efficiency of co-localization between mEGFP-U2AF65 and globin C or globin M-3' in the presence of Phos-stop phosphatase inhibitor</u>	186
<u>5.6 Efficiency of co-localization in control experiments</u>	186
<u>5.7 The effect of SFRS-1-fusion proteins on SFRS-1 3'UTR splicing</u>	190

Introduction

1.1 Spliceosome assembly

1.2 Splice site selection

1.3 Selection of 5' splice sites

1.4 Mechanism of 5' splice site selection

1.5 Selection of 3' splice sites

1.6 Exon definition

1.7 The commitment to splice

1.8 Total internal reflection fluorescence (TIRF) microscopy

Chapter 1. Introduction

Splicing is the process of intron removal from a pre-messenger RNA (pre-mRNA) followed by the subsequent joining of the two exons to form a messenger RNA (mRNA). This process is of high importance as; on average there are 7.8 introns per gene in the human genome (Lander et al 2001).

Chemically, this comprises two S_N2 -type transesterification reactions which are reliant on a series of reactive groups in the intron; these are the guanosine-uridine dinucleotide at the 5' splice site (5' SS), the adenosine at the branch point and the adenosine-guanosine dinucleotide (AG) at the 3' splice site (3' SS; Wahl et al., 2009). These are illustrated in Figure 1.1A. In the first reaction the 2'OH of the branch point adenosine attacks the phosphodiester bond to the 5' side of the 5' SS guanosine; this generates a lariat intermediate and free 5' exon. The 3' OH that is left at the 5' splice site then attacks the phosphodiester bond 3' of the guanosine at the 3' SS, leading to the formation of a lariat product containing the excised intron and an mRNA consisting of the 5' exon joined to the 3'exon (Grabowski et al., 1984; Ruskin et al., 1984; Konarska et al., 1985; Moore and Sharp, 1993; Figure 1.1B).

Remarkably, this is the same mechanism used by Group II self-catalytic introns to splice. Group II introns are found in bacteria, mitochondria and chloroplasts, they have catalytic secondary structures causing them to be excised from the pre-mRNA (Moore and Sharp, 1993; Dayie and Padgett, 2008; Toor et al., 2008). In comparison eukaryotic nuclear introns have little conserved secondary structure due to their divergent nature (Deutsch and Long, 1999) and only have weakly conserved sequences at the 5' SS, 3' SS and branch point.

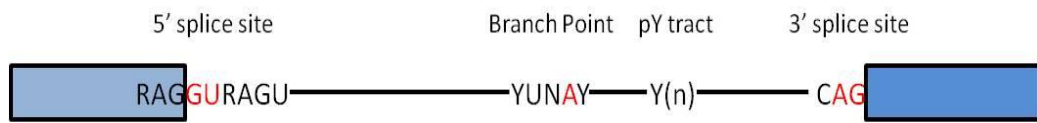
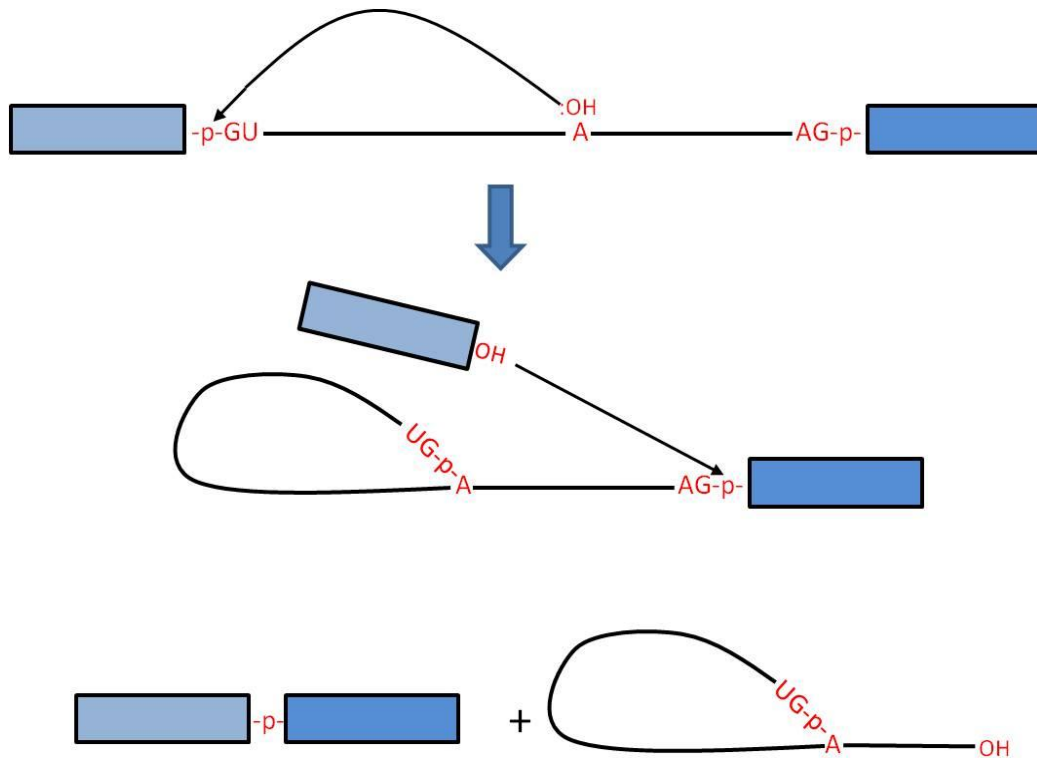
A**B**

Figure 1.1: Pre-mRNA splicing. (A) Conserved sequences within the intron of human genes important for splicing. **(B)** Catalysis of splicing by a two-step transesterification reaction.

Strikingly in many cases Group II introns do not require trans-acting factors to splice. In contrast to this, mammalian introns require many trans-acting factors, which in total are known as the spliceosome (Wahl et al., 2009).

The spliceosome is a highly-regulated biological machine consisting of the U1, U2, U5 and U4/U6 small nuclear ribonucleoparticles (snRNPs) and approximately 170 non-snRNP associated proteins (Wahl et al., 2009). The five snRNPs each contain a small nuclear RNA (snRNA), a set of either seven Sm proteins (U1, U2, U4 and U5) or seven LSM proteins (U6) and a series of snRNP specific factors (Table 1.1). The U4 and U6 snRNPs are extensively base-paired in a di-snRNP and enter the spliceosome as part of a tri-snRNP alongside U5 snRNP (Table 1.1; Will and Luhrmann., 2010). The snRNPs have no active site for catalysis when in their native form; only when the spliceosome as a whole has been assembled can splicing be catalysed.

1.1 Spliceosome assembly

The spliceosome assembles on a pre-mRNA in a series of steps. Studies on *in vitro* spliceosome assembly have been carried out on short introns; the spliceosome assembles across the intron on pre-mRNAs containing small introns of below 200-250nt (Fox-Walsh et al., 2005). In these systems the first specific complex formed is E complex; this is an ATP-independent complex that commits a pre-mRNA to splicing (Reed, 1990; Michaud and Reed, 1991; 1993; Jamison et al., 1992). This complex involves interactions between components of the spliceosome with sequences around both the 5' and 3' SS. Although interactions with one splice site can occur in the

snRNP	U1	U2	U4/U6	U5	U4/U6/U5
Size	12S	17S	13S	20S	25S
snRNA(s)	U1	U2	U4/U6	U5	U4/U6/U5
Proteins	SM 70K A C	SM A' B' SF3a120 SF3a66 SF3a60 SF3b155 SF3b145 SF3b130 SF3b49 SF3b14a/p14 SF3b14b SF3b10	SM LSM hPrp3 hPrp31 hPrp4 CypH 15.5K	SM hPrp8 hBrr2 hSnu114 hPrp6 hPrp28 52K 40K hDib1	SMx2 LSM hPrp8 hBrr2 hSnu114 hPrp6 hPrp28 40K hDib1 hPrp3 hPrp31 hPrp4 CypH 15.5K hSnu66 hSad1 27K

Table 1.1: Protein and RNA composition of the human snRNPs (adapted from Will and Luhrmann, 2010).

absence of the other, they are strongly enhanced on RNAs containing both;
commitment to splicing requires both (Jamison et al., 1992; Michaud and Reed, 1993).

The U1 snRNP and 5' SS recognition in E complex

The first component of the spliceosome to associate with the 5' SS *in vitro* and *in vivo* is the U1 snRNP (Reed, 1990; Wassarman and Steitz, 1992; Gornemann et al., 2005; Huranova et al., 2010). The U1 snRNP contains U1 snRNA, a highly structured RNA (Figure 1.2A; Mount and Steitz, 1981; Branlant et al., 1981) upon which a series of specific proteins are assembled. Seven Sm proteins (D1, D2, F, E, G, D3 and B/B') are assembled around the Sm site (Figure 1.2A) as a heterodimeric ring with each Sm protein binding to a single base in the Sm site consensus (Weber et al., 2010; Raker et

al., 1996; Branlant et al., 1982; Kambach et al., 1999). As well as the core Sm proteins, a series of U1 snRNP-specific proteins associate with the snRNA. The U1 snRNP-specific factor, U1-A associates with stem loop II (Figure 1.2A); U1-A requires no other proteins to bind U1 snRNA and binds through its amino-terminal RNA recognition motif (RRM) (Scherly et al., 1989; Oubridge et al., 1994; Bach et al., 1990). Another U1 snRNP-specific factor, U1-70K associates with the U1 snRNA through its RRM to stem loop I (Patton and Pederson, 1988; Nelissen et al., 1994); it also has contacts elsewhere through its N-terminus which contacts regions around the Sm site (Nelissen et al., 1994; Stark et al., 2001). These contacts are vital for the third U1 snRNP-specific protein, U1-C, which is unable to contact the U1 snRNA in the absence of U1-70K or the Sm proteins (Nelissen et al., 1994). The N-terminus of U1-70K along with the Sm protein D3 creates a binding groove for U1-C, which contacts the minor groove of the duplex between the 5' end of the U1 snRNA (Pomeranz-Krummel et al., 2009) and the 5' splice site (see below). Quantitative mass spectroscopy showed that the U1-associated proteins were present in the snRNP at stoichiometric amounts (Hochleitner et al., 2005).

An association between U1 snRNP and the 5' SS was first suggested based upon the observed complementarity between the 5' end of the U1 snRNA and the 5' SS (Lerner et al., 1980; Rogers and Wall, 1980). When a consensus sequence was derived for the 5' SS it was found to have perfect complementarity to the U1 snRNP (Mount, 1982; Figure 1.2B). The U1 snRNP was subsequently shown to be required for splicing (Padgett et al., 1983; Kramer et al., 1984) and able to bind the 5' SS sequence (Mount

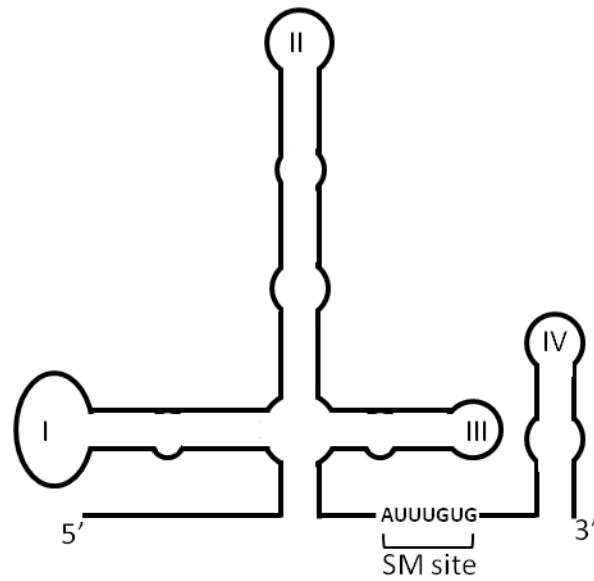
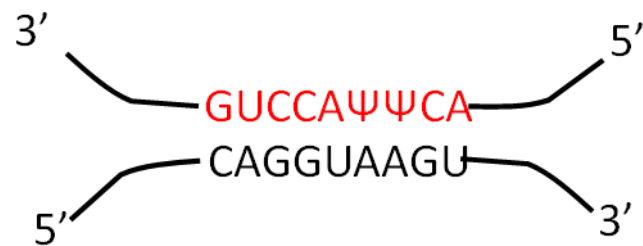
A**B**

Figure 1.2: The U1 snRNP. (A) Structure of U1 snRNA, with stem loops I, II, III, IV and the SM site labelled, re-drawn based on Mount and Steitz, 1981. **(B)** Complementarity between 5' end of U1 snRNA and the consensus 5' splice site, U1 snRNA is in red and pre-mRNA is in black.

et al., 1983) through its 5' terminus (Kramer et al., 1984; Chabot et al., 1985). The importance of this base pairing was illustrated by experiments using U1 snRNAs modified to have higher complementarity to non-consensus 5' SS. Transfection of plasmids encoding these U1 snRNAs led to an increase in the use of splice sites that were previously not recognised by the wild-type U1 snRNP (Zhuang and Weiner, 1986; Zhuang et al., 1987).

The protein component of U1 snRNP was also found to play a role in 5' SS recognition, as partial digestion of U1 snRNP with proteinase K led to a loss of binding (Mount et al., 1983). Work using U1 snRNPs lacking all or some of the U1 snRNP-specific proteins led to the conclusion that U1-C was required for the binding of U1 snRNP to the 5' splice site (Heinrichs et al., 1990), and hence for E complex formation (Will et al., 1996). Truncated U1 snRNP lacking the 5' terminus was subsequently shown to be able to bind the 5' SS consensus sequence under low stringency conditions. This showed that although the 5' terminus of U1 snRNA was needed for stable binding, U1-C was able to select the 5' SS sequence in its absence (Rossi et al., 1996; Lund and Kjiems, 2002; Du and Rosbash, 2002). The importance of U1-C for stable U1 snRNP binding was again demonstrated by mutations in the yeast U1-C homologue that could overcome the requirement for Prp28. Prp28 is a spliceosomal component that is usually needed for the removal of U1 snRNP, and it was inferred that the U1-C mutations had independently facilitated this by weakening U1snRNP binding (Chen et al., 2001).

It is also important to note that in some cases, such as hF1 γ intron 9, U1 snRNP does not seem to be needed for splicing to occur (Fukumura et al., 2009; Crispino et al., 1994).

Recognition of the 3' SS in E complex

The 3' SS region of pre-mRNA is also recognised at an early stage of spliceosomal formation. Recognition of the 3' SS requires the branch point sequence (BPS), the 3' SS AG and a stretch of polypyrimidines inbetween these two motifs (pY tract; Figure 1.1A). In yeast, recognition of this region is dependent upon a strongly conserved BPS and a poor or non-existent pY tract. In contrast, in mammals, recognition is dependent on a poorly conserved BPS and a pY tract. In mammals recognition of the adenosine of the BPS requires a weakly conserved consensus sequence (Gao et al., 2008; Keller and Noon, 1984) with the preferred sequence being UACUAAC (Zhuang et al., 1989; Berglund et al., 1997). The BPS is bound weakly by SF1 (Abovich and Rosbash, 1997; Liu et al., 2001); the interaction of SF1 with the RNA is enhanced by a protein-protein interaction between the amino terminus of SF1 and the third RRM of U2AF65 (Abovich and Rosbash, 1997; Berglund et al., 1998; Gupta et al., 2011; Selenko et al., 2003). U2AF65 is the larger component of a heterodimeric factor U2AF that binds the pY tract (Singh et al., 1995); U2AF65 is a modular protein containing three RRMs and one RS domain (Zamore et al., 1992) and it recognises the pY tract through its first and second RRMs (Zamore et al., 1992; Banerjee et al., 2003; 2004). Binding of SF1 and U2AF65 bends the RNA (Gupta et al., 2011; Kent et al., 2003) allowing the RS domain of U2AF65 to contact the BPS and 3' SS (Valcarcel et al., 1996; Shen and Green, 2004; Kent et al., 2003). Interestingly, substitution of uridines in the pY tract with

pseudo-uridines led to a block in splicing, which was hypothesised to be because of a loss of flexibility (Chen et al., 2010); this suggests that bending of the RNA is very important for the function of U2AF65.

The 3' SS AG is recognised by the smaller component of U2AF, U2AF35 (Wu et al., 1999; Zorio and Blumenthal, 1999; Merendino et al., 1999). The interaction of U2AF35 with U2AF65 (Zamore and Green, 1989), is through a 'proline sandwich' (Kielkopf et al., 2001). This interaction is required for U2AF35 folding (Kellenberger et al., 2002) and has been shown to occur *in vivo* (Rudner et al., 1998; Chusainow et al., 2005). Selection studies using purified U2AF showed that only when U2AF35 is present are pY tracts followed by AG dinucleotides selected (Wu et al., 1999; Singh et al., 1995), showing that U2AF35 enhances the recognition of correct 3' SS.

U2 snRNP association in E complex

The U2 snRNP does not bind the pre-mRNA until A complex (see below), but under low stringency conditions U2 snRNP can be detected as part of E complex (Hong et al., 1997; Jamison and Garcia-Blanco, 1992; Das et al., 2000). E complex cannot form without U2 snRNP, since modified nucleotides in the U2 snRNA have been shown to be needed for E complex formation (Donmez et al., 2004).

Contacts between the splice sites in E complex

Although separate complexes can form on the 5' and 3' portions of a pre-mRNA, E complex formation is enhanced when both splice sites are present (Michaud and Reed, 1993). This has led to the suggestion that there must be contact between the two splice sites in E complex. Experiments to look at the relative position of the 5' and 3'

splice sites (Kent and Macmillan, 2002) and the snRNPs (Donmez et al., 2007) within this complex have shown the two ends of the intron are in close proximity. It has been proposed in yeast that one of the factors making this possible is Prp40, which in yeast is a U1 snRNP factor and can also interact with SF1 bound to the branch point (Abovich and Rosbash, 1997). The mammalian homologue of Prp40, FBP11 has also been implicated in cross-intron contacts (Reed, 2000) and has been found in the mammalian E complex (Das et al., 2000). In another line of evidence it has been proposed that another factor, SC-35 may also play a role in cross-intron bridging due to its ability to interact with the U1 snRNP-specific protein U1-70K and U2AF35 (Wu and Maniatis, 1993). The hypothesis that cross-intron interactions exist between factors at the 5' SS and 3' SS in E complex is also supported by the finding that in extracts depleted of U2AF, an E-like complex (E') can be formed which is dependent on both U1 snRNP and SF1 and in which the 5' SS and BPS are in close proximity (Kent et al., 2005).

A complex assembly

After E complex the first ATP-dependent spliceosomal complex to assemble is A complex (Michaud and Reed, 1991). In this complex the U2 snRNP associates with the BPS (Wu and Manley, 1989) in a U2AF-dependent manner (Ruskin et al., 1988). The interaction of the RS domain of U2AF65 with the BPS enhances association of U2 snRNP (Valcarcel et al., 1996). For U2 snRNP to be able to associate with the BPS the SF1 interaction must be disrupted; it has been hypothesised that the U2AF65 associated RNA dependent helicase, UAP56, may play a role in rearrangements at the BPS resulting in association of the U2 snRNP with the BPS (Fleckner et al., 1997; Kistler and Guthrie, 2001). In yeast, Prp5 has also been implicated in promoting the

interaction between U2 snRNP and the BPS, as its inactivation in yeast extracts led to a loss of U2 snRNP binding (Ruby et al., 1993). A component of the U2 snRNP, SF3 (Table 1.1; Brosi et al., 1993a), is also required for U2 snRNP binding and is needed for A complex formation (Kramer and Utans., 1991). In a biochemical characterization of SF3 it was separated into two components (SF3a and SF3b; Brosi et al., 1993a; b); all of the SF3a/b proteins, apart from SF3b120 were able to crosslink to the pre-mRNA in the spliceosome (Staknis and Reed, 1994a; Gozani et al., 1994). Binding of these proteins in a sequence-independent manner upstream of the BPS was essential for U2 snRNP binding to the BPS and therefore A complex formation (Gozani et al., 1996). An interaction between the SF3 component, SF3b155 and U2AF65 also occurs in A complex (Gozani et al., 1998), this possibly enhances U2 snRNP binding by replacing the interaction between U2AF65 and SF1 that occurs in E complex. Several other A complex-specific factors also associate at this point (Wahl et al., 2009).

Cross-intron interactions in A complex

Further interactions between the 5' SS and 3' SS have also been proposed to occur in A complex. A spliceosomal factor, Prp5, has been implicated in cross-intron interactions at this stage of spliceosome formation as interactions between Prp5 and both of the U1 and U2 snRNPs are required for A complex formation (Xu et al., 2004). It has also been suggested that a co-activator, SR-related matrix protein of 160kD/300kD (SRm160/300kD), also plays a role in cross-intron interactions in a pre-mRNA-specific manner. SRm160/300kD can interact with the spliceosome in an U1 snRNP-dependent manner that is stabilised by U2 snRNP binding (Blencowe et al., 1998). Depletion of U1 snRNP led to the loss of SRm160/300kD and thus a loss of U2 snRNP binding,

suggesting SRm160/300kD plays an important role in cross-intron interactions on a subset of pre-mRNAs (Eldridge et al., 1999).

Spliceosome formation post A complex and catalysis

The formation of B complex is the next stage in spliceosome assembly; B complex is formed from A complex, upon the addition of the U4/U6·U5 tri-snRNP, the Prp19/CDC5 complex, the Retention and Splicing complex and also B complex-specific factors (Wahl et al., 2009). Although B complex is the only complex containing all the snRNPs necessary for splicing it is not catalytically active; catalytic activation requires substantial rearrangement of the components of the spliceosome. During spliceosome activation, U4/U6 snRNP base pairing is destabilised by the DExD/H box protein Brr2 (Will and Luhrmann, 2010); the base pairing between U1 snRNP and the 5' SS of the RNA is destabilised by the DEAD box helicase Prp28p and replaced by base pairing between the 5' SS and U6 snRNP (Staley and Guthrie, 1999; Chen et al., 2001). These rearrangements, along with a substantial alteration in the protein composition of the spliceosome, give rise to B^{ACT} complex (Bessonov et al., 2010). B^{ACT} complex is then structurally remodelled by Prp2 to form B* complex (Warkocki et al., 2009). The first transesterification reaction of splicing can then occur, with the bulged adenosine nucleotide in the BPS (Query et al., 1994) attacking the phosphodiester bond at the 5' SS. This leads to the formation of C complex, the final complex in spliceosome assembly. C complex formation requires further compositional remodelling whereby several C complex and step-II-specific factors associate (Wahl et al., 2009). For the second transesterification reaction to occur, substantial rearrangement of the snRNP-RNA interactions is required to form the second catalytic centre of the spliceosome

(Konarska et al., 2006). Once the second transesterification reaction has occurred, the spliceosome is disassembled, releasing the mRNA and intron lariat. The spliceosomal factors are then recycled allowing them to partake in further rounds of splicing (Wahl et al., 2009). The later stages of spliceosome assembly are a highly regulated, yet dynamic series of events which are controlled by many regulatory protein factors including DExD/H box proteins and possibly peptidyl-prolyl cis/trans isomerases (Wahl et al., 2009).

1.2 Splice site selection

The brief account of the spliceosomal assembly pathway described in section 1.1 is based on experiments with pre-mRNA substrates containing strong splicing signals, but the consensus splice site signals rarely exist in nature (Mount, 1982). The spliceosome is able to recognise sequences quite unlike the consensus. To enable recognition of the non-consensus sequences the spliceosome requires other trans-acting factors to bind the pre-mRNA and enhance the detection of splice sites (Xiao et al., 2007). These factors bind sequences generally in exons, but also in introns, known as splicing enhancers. The first splicing enhancer was identified in the fibronectin EDIII exon, which, failed to splice when part of the exon was inverted (Mardon et al., 1987). These sequences are bound by trans-acting factors that are able to enhance the recognition of splice site signals and progression through spliceosomal assembly (see below). The best defined set of enhancer trans-acting factors are the SR proteins; these proteins have RRM domains (usually one or two) and RS domains that are rich in arginine and serine residues (Graveley, 2000; Long and Cáceres, 2009).

As well as splicing enhancer sequences, there are splicing inhibitory sequences. These are bound by negative regulators of splicing and splice site recognition such as members of the hnRNP family. Although, in general, SR proteins are enhancers and hnRNP proteins are repressors, a large body of evidence suggests they can play both roles in a context-dependent manner (Martinez-Contreras et al., 2007; see below). Factors that bind these two types of elements are needed to regulate splice site selection; they work in a combinatorial fashion to ensure correct splice site selection (Ben-Dov et al., 2008; Matlin et al., 2005; Smith and Valcarcel, 2000) and can even compete to bind the same sequences (for example see Rooke et al., 2003; Zahler et al., 2004). This has led to the suggestion that there may be a “splicing code”, knowledge of which would allow patterns of splicing to be predicted (Barash et al., 2010). As splice site sequences resemble the consensus sequence less, the more they rely upon splicing enhancers to be recognised (Xiao et al., 2007). This in turn allows alternative patterns of splicing to develop in cell-(Black, 1998) and tissue-(Xu et al., 2002; Wang et al., 2008; Castle et al., 2008) specific manners, as well as in response to external stimuli (Stamm, 2002). Therefore the spliceosome is a dynamic and adaptable machine and the pattern of regulation is critical for many aspects of life at a whole organism level; this is exemplified by the finding that in humans up to 60 % of disease-causing mutations are predicted to affect splicing (Lopez-Bigas et al., 2005).

1.3 Selection of 5' splice sites

When an intron unit contains more than one 5' splice site the spliceosome must make a decision as to which one is used. Failure to make the correct decision will lead to the production of aberrantly spliced transcripts. Diseases such as β -thalassemia can be

caused by mutations in the 5' SS sequence leading to cryptic 5' SS usage (Treisman et al., 1983).

The affinity effect

The overall affinity of U1 snRNPs for a 5' splice site has been shown to be very important in splice site use. Base pairing between the U1 snRNA and 5' SS clearly plays an important role in selection, as transfection of U1 snRNA with modified sequences at the 5' terminus into cells can alter 5' SS choice (Zhuang and Weiner, 1986; Zhuang et al., 1987; Yuo and Weiner; 1989, Freund et al., 2003; Sorek et al., 2004). Also, when different 5' SS sequences were tested in competition assays, those sequences with a high predicted affinity for U1 snRNP were favoured (Eperon et al., 1986; Lear et al., 1990; Zhou et al., 2008).

Despite the favouring of strong 5' SS, and the selection against weak 5' SS, the correlation between 5' SS use and predicted U1 snRNP base-pairing breaks down in cases where base-pairing is predicted to be moderate (Lear et al., 1990; Roca et al., 2005). Although this in part may be explained by shifted base-pairing between the U1 snRNP and 5' SS (Roca and Krainer, 2009), several early observations show other sequences must be involved. First, in nature the strongest predicted sites are not always used (Shapiro and Senapathy, 1987), second, splice site use does not always correlate with U1 snRNP binding (Mayeda and Ohshima, 1988; Nelson and Green, 1988; Nelson and Green, 1990; Eperon et al., 1993) third, tandem duplications of 5' SS need exonic sequences to be spliced (Reed and Maniatis, 1986). Exonic sequences are needed for efficient use of 5' SS because of the trans-acting factors that bind within the exons.

Many trans-acting factors have been found to be able to influence 5' splice site selection through binding to regulatory sequences. Through *in vitro* analysis it has been shown that SR proteins (such as SFRS-1 and SFRS-2; previously known as SF2/ASF and SC-35 respectively) are able to enhance the binding of U1 snRNP to the RNA (Eperon et al., 1993). Both SFRS-1 and SFRS-2 have been shown to interact with U1-70K (Wu and Maniatis, 1993; Kohtz et al., 1994; Cho et al., 2011) in a phosphorylation-dependent manner (Xiao and Manley, 1997; Cho et al., 2011) either via the RS domains (Wu and Maniatis, 1993; Kohtz et al., 1994) or through the RRM of the SR protein (Cho et al., 2011). SFRS-1 also binds co-operatively with the U1 snRNP (Kohtz et al., 1994) and possibly contacts the 5' SS sequence directly (Zuo and Manley, 1994). This effect appears to be complementary, as U1 snRNP has been shown to stabilise the binding of SFRS-1 to RNA (Jamison et al., 1995). Experiments where the effect of SR proteins and the interaction between SFRS1 and U1-70K on E complex formation were investigated, suggest that this effect is not an artefact, as both are needed for E complex formation in a 5' SS-dependent manner (Staknis and Reed, 1994b; Cho et al., 2011).

In general, SR proteins have been shown to shift splicing to downstream 5'splice sites in competition assays based on alternative 5' SS both *in vivo* (Misteli et al., 1997; Caceres et al., 1997; Wang et al., 1998) and *in vitro* (Ge and Manley, 1990; Krainer et al., 1990; Caceres and Krainer, 1993; Eperon et al., 1993; Screatton et al., 1995; Eperon et al., 2000) and to induce exon inclusion (Cote et al., 1999; Selvakumar and Helfman, 1999). This effect can be modulated, at least in the case of SFRS-1, by increasing the binding of U1 snRNP to the RNA (Eperon et al., 1993, Eperon et al., 2000) and is not

dependent on the RS domain of the SR protein (Caceres and Krainer, 1993; Eperon et al., 2000; Caceres et al., 1997). Exceptions to this general statement include SFRS11 (p54) which is able to shift splicing to the upstream 5' SS (Zhang and Wu, 1996).

The SR proteins are also important in 5' SS definition later in spliceosome assembly; they have been implicated in the recruitment of the U4/U6·U5 tri-snRNP to the spliceosome (Rosciigno and Garcia-Blanco, 1995). The RS domains of SR proteins also contact the 5' SS in B complex, stimulating the binding of U6 snRNP to the 5' SS (Shen and Green, 2004; 2006). The RS domain continues to contact the 5' SS into the later stages of spliceosome formation and stimulates the binding of U5 snRNP to the region upstream of the 5' SS (Shen and Green, 2007). The ability of SR proteins to recruit the U6 snRNP may explain why on some RNA transcripts, large amounts of SR proteins can compensate for a loss of U1 snRNP (Tarn and Steitz., 1994; Crispino et al., 1994). Interestingly, interactions between the U5 snRNP (Cortes et al., 1993) or U6 snRNP (Hwang and Cohen, 1996) and pre-mRNA have been shown to affect 5' SS selection, showing that there are multiple mechanisms for 5' SS selection.

Binding of U1 snRNP to the 5' splice site is influenced by many other proteins as well as the SR proteins. For example, Tia-1 binds U-tracts downstream of the 5' SS and stimulates U1 snRNP binding (Forch et al., 2000) through an interaction with U1-C (Forch et al., 2002); this can activate the 5' SS causing it to be used (Forch et al., 2000 Gal-Mark et al., 2009; Wang et al., 2010). Interestingly the enhancing effect of Tia-1 can be counteracted by sex lethal (SXL) in *Drosophila melanogaster* (Forch et al., 2001) and polypyrimidine tract binding protein (PTB) in Humans (Shukla et al., 2005), which compete with Tia-1 for binding to the uridine-tract. A protein normally associated

with the 3' SS, U2AF65, can also bind these U-tracts and stimulate the binding of U1 snRNP to the 5' SS (Forch et al., 2003). Another example of a splicing enhancer is RBM25, which has been shown to enhance the stability of U1 snRNP binding to 5' SS and thus influence 5' SS selection (Zhou et al., 2008).

Splicing inhibitors can also influence 5' SS selection and the binding affinity of U1 snRNP (Yu et al., 2008). One such splicing inhibitor, hnRNP A1, a well-characterised splicing regulator, has been shown to influence the pattern of 5' SS selection. It increases the level of upstream 5' SS use in competition assays (Yang et al., 1994; Mayeda et al., 1994; Shen et al., 1995; Chabot et al., 1997) and is thought to do this by either competing with U1 snRNP for binding to the 5' SS (Eperon et al., 2000) or by looping out sections of the pre-mRNA (Blanchette and Chabot, 1999). In 5' SS competition and exon skipping assays, hnRNP A1 has been shown to antagonise the effect of SFRS-1 (Mayeda and Krainer, 1992; Mayeda et al., 1993; Sun et al., 1993; Caceres et al., 1994; Eperon et al., 2000; Zhu et al., 2001) and it is thought that SFRS-1 blocks binding of hnRNP A1 to the RNA (Eperon et al., 2000; Zhu et al., 2001; Sun et al., 1993). This interplay between SFRS-1 and hnRNP A1 is backed up by the finding that the ratio between them can vary *in vivo* in a spatial and temporal manner (Pollard et al., 2000).

Another member of the hnRNP protein family, hnRNP H, has also been shown to be important in 5' SS selection. Its action is complex, as in some cases it is able to compete with U1 snRNP in binding to the 5' SS and inhibit the use of the splice site (Buratti et al., 2004), whilst in other cases it co-operates with SR proteins and acts as an enhancer of U1 snRNP recruitment (Caputi and Zahler, 2002). Members of the

hnRNP protein family can also act together to modulate selection of 5' splice sites by looping out sections of the RNA (Fisette et al., 2010), thus influencing splice site pairing, and showing that in different contexts different proteins can have different effects.

The affinity of U1 snRNP for a 5' SS is also modulated by the availability of the 5' SS for base-pairing. The secondary-structure landscape of an RNA can modulate 5' SS use; when a 5' SS is sequestered in native (Blanchette and Chabot, 1997; Abbink and Berkhout, 2008; Singh et al., 1997; Loeb et al., 2002; Jiang et al., 2000), or artificially engineered (Solnick and Lee, 1987; Eperon et al., 1988) secondary structures the U1 snRNP cannot recognise the 5' SS and it is not used.

The position effect

All of the described regulatory factors, along with many others are able to influence 5' SS selection. However, affinity alone cannot explain all the observed 5' SS selection patterns. There is a well-documented position-dependent effect on 5' SS selection, whereby when two 5' SS of comparable strength are competing, there is a strong preference for the downstream 5' SS (Eperon et al., 1993; Cunningham et al., 1991; Hicks et al., 2010; Reed and Maniatis., 1986; Yu et al., 2008). This preference has been observed, in some cases, even if the downstream 5' SS is predicted to have a lower ability to bind U1 snRNP (Hicks et al., 2010). The position effect is also dependent on the separation between the two splice sites. In constructs where the splice sites are separated by less than 40 nucleotides, both splice sites are used, but when they are separated by more than 40 nucleotides the downstream 5' SS is highly favoured (Cunningham et al., 1991).

Selection of 5' splice sites: post U1 snRNP binding

Splicing inhibitors can also inhibit 5' SS selection after U1 snRNP binding: for example, HMGA1a can trap a U1 snRNP bound to a 5' splice site causing it not to be recognised and the exon to be skipped (Ohe and Mayeda, 2010); PSI can stimulate U1 snRNP to bind to pseudo-5' splice sites possibly causing steric occlusion of the true 5' SS (Labourier et al., 2001). Other inhibitors, such as PTB and RBM-5 have been shown to inhibit the association between the 5' SS and downstream 3' SS thus inhibiting splicing (Sharma et al., 2008; Bonnal et al., 2008; Sharma et al., 2011), with PTB possibly binding to stem loop IV of the U1 snRNA and inhibiting spliceosome formation (Sharma et al., 2011).

1.4 Mechanism of 5' splice site selection

The selection of 5' splice sites was originally proposed to be due to the affinity of the U1 snRNP for the 5' SS assuming that only one snRNP may be present on a single RNA. The number of U1 snRNPs bound was hypothesised to be restricted to one by either a kinetic model, where as soon as a U1 snRNP bound, it would be paired with components at the 3' SS and thus used, or by a sequestration model, wherein the single U1 snRNP would be recruited first by factors at another site on the RNA such as the cap binding complex or components bound to the 3' SS (Figure 1.3; Lewis et al., 1996; Cunningham et al., 1990; Eperon et al., 1986).

However, predictions made based on these models did not agree with two other observations, firstly, the dissociation constant for a strong splice site made it improbable that a strong splice site would not be bound by U1 snRNP (Eperon et al.,

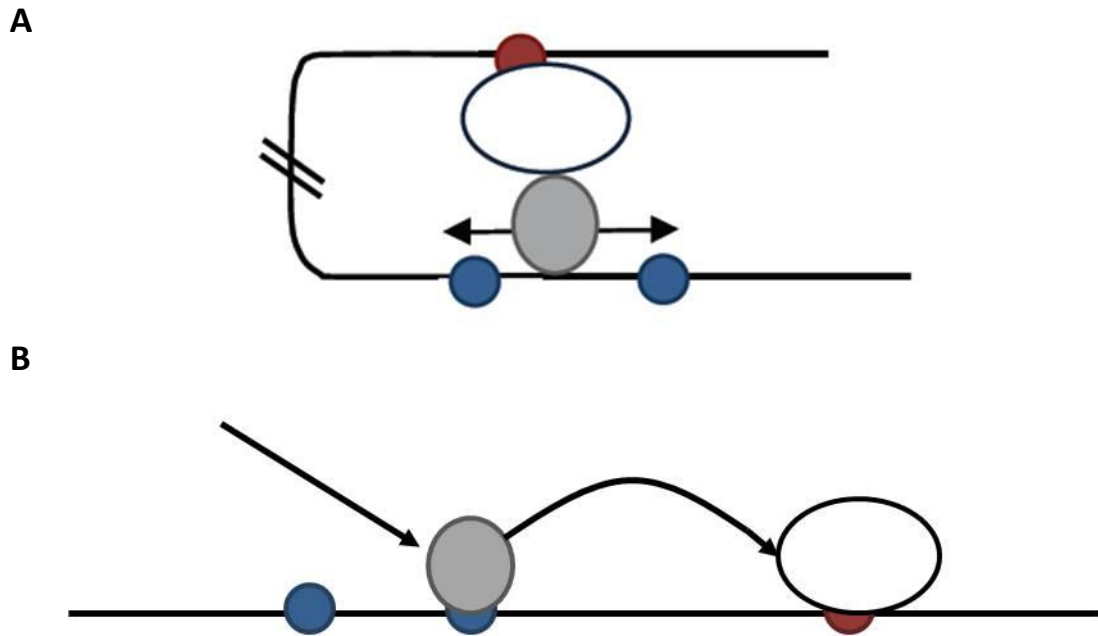


Figure 1.3: Mechanisms of 5' splice site selection based upon affinity. (A) Shows a mechanism where U1 snRNP is recruited first by other factors and then possible 5'ss are sampled. **(B)** Shows an alternative model based upon kinetics whereby the first U1 snRNP bound is selected.

2000); secondly the models based solely on affinity could not be used to explain the position effect (see above). This is because if there are two 5' SS of equal strength, then either of the affinity models would predict equal use of the two 5' splice sites.

If, however, the selection was not based solely on the affinity of a single U1 snRNP, then the observed position effect could be explained by a model where both 5' splice sites are occupied by U1 snRNP independently. In this model, the independent occupation hypothesis, an unknown process would lead to the downstream 5' SS being selected in preference to the upstream 5' SS. The independent occupation hypothesis can also incorporate the effects of affinity on 5' SS selection, since position will only have an effect when the two 5' SS are both occupied. In cases where the 5' SS are of different affinity towards the U1 snRNP, the percentage occupation of the two 5' SS will play a part in 5' SS selection. Different proportions of the pre-mRNA population will be occupied at either site, or both sites, leading to different usages of the two 5' SS dependent on how many of the splice sites are occupied by U1 snRNP (Figure 1.4). This is supported by the finding that in constructs where the downstream 5' SS is weaker than the upstream 5' SS the downstream 5' SS can still be preferred (Hicks et al., 2010).

The independent occupation hypothesis also provides an explanation for the observed effects of SR proteins and hnRNP A1 on 5' SS selection. In general SR proteins shift splicing towards the downstream 5' SS in competition assays (Section 1.3) and they increase binding of U1 snRNP to both sites (Eperon et al., 2000). This can be explained by the position effect where having both sites occupied will lead to the downstream 5'

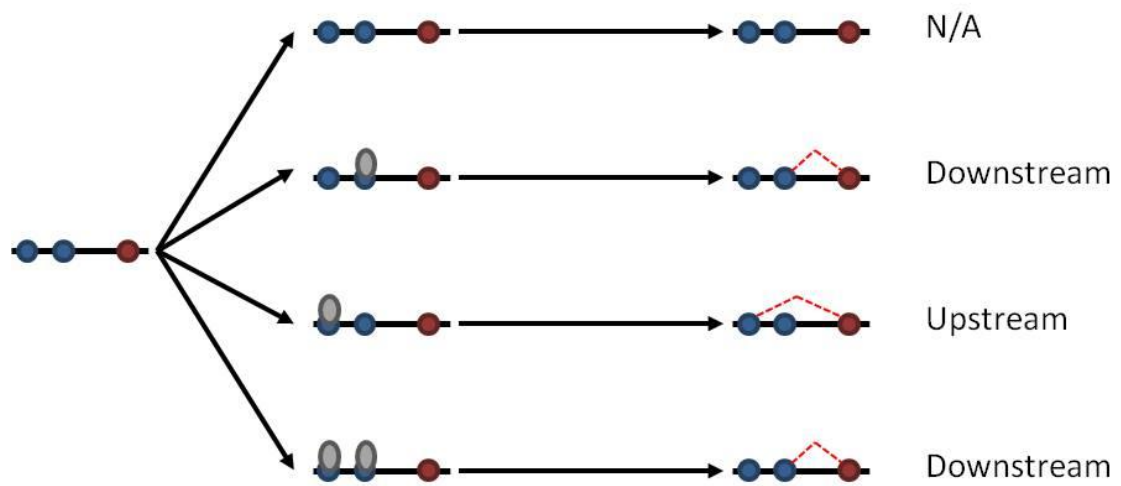


Figure 1.4: The independent occupation model. Illustration showing the possible splicing outcomes which would result from different occupancy of different 5' SS by U1 snRNPs.

SS being used. In contrast hnRNP A1 shifts splicing towards the upstream 5' SS (Section 1.3) and has been observed to cause a reduction in the binding of U1 snRNP to the downstream 5' SS (Eperon et al., 2000). Therefore, the proportion of the RNA population occupied at both sites is reduced and thus the upstream site is used.

Direct evidence for the presence of two U1 snRNPs on a single RNA is lacking; U1 snRNP-dependent protection of two 5' SS in an RNA has been detected in RNase protection assays (Eperon et al., 1993; Nelson and Green, 1988). This data, however, does not provide direct evidence that two U1 snRNPs are bound and could be explained by the protection of the other 5' SS by other factors in a way that is dependent on U1 snRNP being bound to the other 5' SS. For example, both SFRS-1 and hnRNP H have been shown to bind 5' SS directly and the interaction between SFRS-1 and pre-mRNA has been shown to be stabilised by U1 snRNP (Section 1.3). It is also possible that the observed protection is due to proteins binding to the exon as part of exon definition (Section 1.6).

The hypothesis of independent occupation has been used in recent studies to explain observations such as an upstream 5' SS being able to enhance the use of a downstream 5' SS (Hicks et al., 2010) and the effect of silencer sequences on 5' SS selection (Yu et al., 2008). The observed position effect has been hypothesised to be due to the random diffusion of the RNA, meaning that those 5' SS that are closer to the 3' SS are more likely to encounter the 3' SS and therefore to be selected for splicing (Yu et al., 2008). However, this has never been tested and other explanations are possible.

1.5 Selection of 3' splice sites

Recognition of the 3' SS is mainly determined by recognition of the pY tract, as deletion of the pY tract leads to complete loss of splicing (Reed and Maniatis, 1985; Reed et al., 1989; Zhu and Krainer, 2000). In contrast the BPS, although important in increasing binding of U2AF65 to the pY tract (Berglund et al., 1998; Gupta et al., 2010), is less important in 3' SS selection. Deletion of the normal BPS will lead to activation of cryptic branch point sequences (Ruskin et al., 1985) and, despite there being a preferred BPS (Zhuang et al., 1989), in some cases multiple branch point sequences can be utilised in a single intron (Noble et al., 1987; Helfman et al., 1989). The BPS cannot be recognised on its own as moving a BPS away from the adjacent pY tract leads to activation of cryptic branch point sequences instead of recognition of the original (Ruskin et al., 1985). The AG at the 3' SS is very important for the second transesterification reaction of splicing, but its importance in the first step of splicing is dependent on the strength of the pY tract (see below).

Recognition of the polypyrimidine tract

The pY tract is first recognised by the large subunit of U2AF (U2AF65), which recognises sequences that are highly enriched in uridine (Singh et al., 1995). Only the first and second RRM bind specifically (Banerjee et al., 2003; 2004). Despite the preference for uridine, there is wide variation in the composition of natural pY tracts (Mount, 1982; Shapiro and Senapathy, 1987). Mutational analysis of pY tracts has shown that in general stronger pY tracts have a higher content of uridine although pY tracts with lower uridine contents are regularly used (Roscinio et al., 1993; Coolidge et al., 1997). The length of the pY tract is also important, with longer pY tracts being

more efficiently recognised (Reed, 1989). A short pY tract often requires a strong BPS to be recognised (Reed, 1989; Smith et al., 1989; Mullen et al., 1991; Coolidge et al., 1997), presumably due to the stabilising effect of SF1 (see above), whilst a strong pY tract is recognised independently of the BPS.

Experiments by Banerjee et al. (2003) led to the hypothesis that U2AF65 was able to cope with this natural variation in pY tract composition and length because it had multiple modes of binding. They hypothesised that U2AF65 can bind pY tracts with relatively low levels of uridine and different lengths, because it can loop out bases that are not efficiently bound (Banerjee et al., 2003). This suggestion was backed up by the finding that U2AF65 recognises bases based on the hydrogen bonds they form with amino acid side chains within U2AF65, which can be “flipped out” of the way when necessary, thus showing how U2AF65 could tolerate bases other than uridine (Sickmier et al., 2006). Further analysis of U2AF65 binding also showed that the two RRM s can bind with relative independence. This was thought to be due to the flexible linker between the two RRM s possibly allowing them to sample until the best site is found (Jenkins et al., 2008). Banerjee et al. (2004) later published experiments where sequences with high affinity for U2AF65 were selected and using electrophoretic mobility shift assays the ability of U2AF65 to bind them was assessed. Close inspection of this data suggests another possible explanation for the effect of pY tract length on splicing; four out of the six sequences tested had multiple higher-order complexes that appeared in a U2AF65 concentration-dependent manner, suggesting that longer pY tracts may be able to bind more than one U2AF65.

The role of regulatory factors In 3' splice site selection

Those pre-mRNAs with pY tracts that are weak, either because they have low levels of uridine or are short, often depend on other sequences to enable the first transesterification reaction to occur (Wang et al., 1995). In these cases, the AG dinucleotide at the 3' SS and splicing enhancer sequences have been shown to be of great importance. The roles of these features, in 3' SS selection have been the subject of many studies.

Firstly, it was shown that SR proteins were able to interact with the RS domain of U2AF35 but not U2AF65 in vitro (Wu and Maniatis, 1993). Based on the interaction between U2AF65 and U2AF35 (Zamore et al., 1989; Keilkopf et al., 2001) it was then suggested that the dependency on the 3' SS and enhancer sequences was because SR proteins were able to recruit U2AF35 to the 3' SS and in doing so were able to recruit U2AF65 to the pY tract (Tian and Maniatis, 1993; Tian and Maniatis, 1994; Zuo and Maniatis, 1996; Zhu and Krainer, 2000; Henscheid et al., 2008). This was supported by cross-linking studies showing that U2AF65 binding to weak 3' splice sites was dependent on enhancer sequences and U2AF35 (Zuo and Maniatis, 1996; Wu et al., 1999; Graveley et al., 2001; Wang et al., 1995; Bouck et al., 1998) and by studies showing that RNAs with long pY tracts could be rendered enhancer-dependent by shortening the pY (Tian and Maniatis, 1992; Graveley and Maniatis, 1998; Zhu and Krainer, 2000).

In contrast, other experiments on the same pre-mRNAs have shown that instead of enhancing U2AF65 binding, SR proteins function to block splicing inhibitors bound elsewhere on the pre-mRNA (Kan and Green, 1999). Experiments looking at the effect

of increasing concentrations of U2AF65 in the absence of U2AF35 showed that although U2AF65 binding could be stimulated, splicing still did not occur. This led to the hypothesis that although U2AF35 can stimulate U2AF65 binding, it is not the sole function of U2AF35 in the early stages of splicing (Guth et al., 2001).

A separate set of experiments has suggested that some splicing enhancers may be able to enhance the recruitment of U2 snRNP to the BPS in a position-dependent manner, instead of recruiting U2AF65. Mutation of a splicing enhancer in exon 6 of the β -tropomyosin gene led to stalling of splicing prior to A complex formation, consistent with an effect on U2 recruitment (Selvakumar and Helfman, 1999). In addition, some studies have shown that enhancer sequences do not stimulate increased U2AF65 binding to the pre-mRNA (Li and Blencowe, 1999; Guth et al, 1999), suggesting that regulation occurs after recognition of the pY tract by U2AF65. Through a series of elegant experiments, it has been proposed that the RS domains of SR proteins are able to enhance splicing by contacting the BPS. This contact was found to be ATP and U2 snRNA dependent, which suggests that it occurs in A complex (Shen et al., 2004, Shen and Green, 2004). In further experiments they went on to show that the interaction of the RS domain with the BPS requires double stranded RNA and that increasing concentrations of an RS domain-containing protein stimulated U2 snRNA binding (Shen and Green, 2006). A role for SR proteins in U2 recruitment is supported by the observations that U2 snRNP binding does not always correlate with U2AF65 binding and that inhibition of an enhancer in exon 7 of SMN1/2 can block U2 snRNP binding but not U2AF65 binding (Martins de Araujo et al., 2009).

Although these two models for enhancer function are not necessarily mutually exclusive, the fact that contradictory conclusions have been drawn for the same pre-mRNAs under different experimental conditions make it difficult to accept both (see Kan and Green, 1999; Gravely et al., 2001; Li and Blencowe, 1999).

The effect of SR proteins on U2 snRNP recruitment is position-dependent. Intronic sequences that bind SR proteins have been found to repress the use of 3' splice sites and U2 snRNP recruitment (Simard and Chabot, 2002, Kanopka et al., 1996) and in 3' SS competition assays, exonic sequences upstream of a 3' SS led to inhibition of splicing even if the competing 3' SS was deleted (Ibrahim et al., 2005). Interestingly, moving the inhibitory sequence into an exon led to it becoming an enhancer (Kanopka et al., 1996), showing that SR proteins will have different effects dependent on their position within the pre-mRNA.

The selection of 3' splice sites in competition experiments is affected by SR proteins. As with their general effect on 5' SS selection they have been found to enhance the use of the proximal 3' SS (Fu et al., 1992; Bai et al., 1999); this effect has also been shown to be countered by hnRNPA1 (Bai et al., 1999) as is the case for 5' SS (Section 1.3).

Other inhibitors of 3' SS selection have also been studied in great detail. For example, SXL is a pY tract-binding protein (Singh et al., 1995) which is very important in determining the sex of *Drosophila melanogaster*. The selection of 3' SS is affected by SXL in two ways: firstly it can directly compete with U2AF for binding of the pY tract as in the case of the transformer and MSL-2 RNAs (Sosnowski et al., 1989; Inoue et al., 1990; Valcarcel et al., 1993; Merendino et al., 1999); secondly, it is able to influence

selection at the second step of splicing, as is the case for autoregulation of its own RNA where SXL binds to the U2 snRNP factor SPF45, inhibiting the second step of splicing and leading to skipping of exon 3 (Lallena et al., 2002; Chaouki and Salz., 2006). It has also been suggested that PTB is able to compete with U2AF65 for the binding of pY tracts on some RNAs (Sauliere et al., 2006). However, recent experiments suggest that at least in cases where there are multiple binding sites, it is likely that PTB functions by inhibiting recognition of the whole exon by looping out the exon (Oberstrass et al., 2005; Cherny et al., 2010). It has also been hypothesised that some inhibitory proteins, such as hnRNP H, can bind to sequences near the 3' SS and inhibit recognition by directly inhibiting U2AF35 binding (Jacquenet et al., 2001; Romano et al., 2002).

Selection of the AG di-nucleotide

After U2 snRNP binding but prior to the second transesterification reaction, a second selection step occurs. The AG di-nucleotide that will be used in the second step of splicing must be selected. Selection of the correct AG was originally proposed to be by a scanning mechanism where the first AG detected downstream of the BPS is used. This was based on the observation that insertion of an AG upstream of the natural AG led to activation of the inserted AG. The level of activation of the inserted AG was determined by the preceding nucleotide, with CAG, AAG and TAG all being used efficiently but not GAG (Smith et al., 1989).

This hypothesis was extended to include an AG “exclusion zone”, whereby an AG that is too close to the BPS is not used and instead an AG further downstream is used. It was suggested that the reason for this was steric hindrance caused by factors bound at

the BPS (Smith et al., 1989; 1993; Chua and Reed, 1999); subsequent experiments have shown many introns to have AG exclusion zones (Gooding et al., 2006). The scanning model was also supported by bimolecular ligation assays which showed that the first AG was used even when the BPS and 3' SS were not connected. This selection required a free 5' end for the portion of the RNA containing the 3' SS, suggesting that the spliceosome scans from the 5' end to the nearest AG (Anderson and Moore., 1997, Chen et al., 2000). Experiments where 3' SS AGs were sequestered in hairpin structures also showed that they could be bypassed, as would be predicted by the scanning model (Smith et al., 1993). This model was also used to explain why some introns had branch point sequences 170nt upstream of the AG when the average distance was much lower. In these cases the AG used for splicing was the first AG downstream of the BPS (Smith and Nadal-Ginard, 1989; Helfman and Ricci, 1989).

The discovery that two AG dinucleotides in close proximity to one another could compete led to an adaptation of the scanning model (Smith et al., 1993; Chua and Reed, 2001). The competition was dependent on the position of the AGs relative to both the branch point and each other, meaning that the first AG was not always the one used (Smith et al., 1993; Chua and Reed, 2001). This led to the proposal that scanning and competition were both important (Smith et al., 1993) so that if two AGs were in close proximity they could compete for use whilst if they were separated by a long stretch the first AG was used.

In some cases, very close AG dinucleotides appear to be very important; thirty percent of human RNAs have been found to have NAGNAG 3' SS sequences, where N can be any base (Hiller et al., 2004). It has also been estimated that half of all introns that

show alternative 3' SS use involve NAGNAG sequences (Akerman and Mandel-Gutfreund, 2006). In exonised Alu elements the presence of an NAGNAG 3' SS has been shown to be important in ensuring that the exons remain non-constitutive, presumably because the two AG sequences interfere with correct recognition of the 3' SS (Lev-Maor et al., 2003). It has also been suggested that the subtle alteration in protein sequence allowed by a NAGNAG 3' SS is of importance for some proteins (Tsai et al., 2008). Interestingly, this step in 3' SS selection can be regulated. Protein $G\alpha_s$ splicing involves an unusual NTGNAG 3' SS; SFRS-1 promotes the use of the TG splice site whilst hnRNPA1 promotes the use of the AG 3' SS (Pollard et al., 2002).

Both hSlu7 and the U5 snRNP have also been implicated in the process of AG selection; hSlu7 has been shown to be important as its deletion led to the use of AG sites that were previously too close to the BPS to be used. It was suggested that this was because hSlu7 held the 5' exon in the spliceosome in a certain position and when it was deleted, the 5' exon it was not as stable and therefore was not necessarily in the correct orientation to attack the selected AG (Chua and Reed, 1999).

The U5 snRNP has been proposed to play a role in 3' SS selection based on the identification of U5 snRNP components that were cross-linked to the pY tract in C complex. Interestingly, when an AG was too close to the BPS to be used cross-links were detected between the hSnu114 protein and the 3' SS, suggesting steric hindrance. The U5 snRNP proteins appeared to require a pY tract in order to bind to the RNA and a pY tract close to the AG increased the kinetics of the second transesterification reaction (Chiara et al., 1997). This was in agreement with some work suggesting that a pY tract close to the AG was advantageous (Reed et al., 1989;

Frederick and Keller, 1985) but contradicting other work showing the opposite (Smith et al., 1989).

Overall, from the examples presented for 5' SS and 3' SS selection, it is clear that there are multiple ways in which regulation can be achieved. The spliceosome appears to be an incredibly dynamic machine, able to adapt to influences from a wide range of factors.

1.6 Exon definition

Since enhancers are able to enhance recognition of both the 5' and 3' splice sites it follows that enhancers would be able to stimulate both at once through a series of protein-protein interactions (Wu and Maniatis, 1993; Bourgeois et al., 1999; Lam and Hertel, 2002). This idea of a connection between the 5' SS and 3' SS during splice site recognition is called Exon Definition and was originally postulated as a reason for the phenotypes associated with 5' SS deletion. When a 5' SS was deleted it led to the inhibition of the splicing of the upstream intron (Robberson et al., 1990). Exon definition only occurred if the exon was below 300nt long and had 3' and 5' splice sites (Robberson et al., 1990, Berget, 1995). The exon definition model gave an explanation as to how the spliceosome was able to recognise exons when they were separated by thousands of nucleotides of intronic RNA. In accordance with the model, it has been found that mammalian exons average just 120nt in length, which is easily inside the exon definition limit of 300nt (Ast, 2004). The exon definition model is supported by the findings that firstly, deletion of a 5' SS in an internal exon leads to exon skipping (Talerico and Berget, 1990) and secondly, recruitment of the U1 snRNP has been reported to enhance recognition of the 3' SS by U2AF65 and U2 snRNP in a SR protein-

dependent manner (Hoffman and Grabowski, 1992; Boukis et al., 2004). This need for exonic sequences is why when 5' SS are inserted into intronic sequences, they are not used (Nelson and Green, 1988); this may be because the use of a 5' SS is also dependent on the context. The ability to make cross-exon interactions is also a point of regulation, as PTB is able to bind in an exon close to the 5' SS and inhibit exon definition after U1 snRNP recruitment (Izquierdo et al., 2005).

The exact point at which the exon-defined complex is disrupted and replaced with an intron-defined complex is unknown. The existence of interactions between the two splice sites when short introns are used (Fox-Walsh et al., 2005) points to the idea that complexes can form across introns at very early stages. Experiments have shown that PTB, RBM5 and hnRNP L regulate splicing by blocking the transition from exon-defined to intron-defined complexes after the binding of U1 and U2 snRNPs. This suggests that they block the transition after A complex formation (House and Lynch, 2006; Bonnal et al., 2008; Sharma et al., 2008; 2011). This has led to the hypothesis that the switch often occurs in A complex, although the recent discovery that the U4/U6·U5 tri-snRNP can associate with exon-defined complexes (Schneider et al., 2010) again complicates things. This suggests that there are many points at which the switch can occur and that it may be dependent on the signal contained within a specific RNA. Interestingly, the hnRNP proteins have been implicated in enhancing the pairing of splice sites across an intron by looping segments of the intron out. The hnRNPs bind either end of an intron and bring the 5' and 3' splice sites closer together (Martinez-Contreras et al., 2006). Whether splice sites are defined by exon definition or intron definition has also been reported to play a role in whether or not they are used in splicing, suggesting

that the length of the introns surrounding an exon play an important role in splice site selection (Fox-Walsh et al., 2005).

1.7 The commitment to splice

An RNA is committed to being spliced once E complex is formed. This was shown by experiments in which purified E complex was chased into later splicing complexes (Michaud and Reed, 1991). At this early stage in spliceosome formation, the two splice sites and components bound to them have been shown to be close together in constructs with a single 5' SS and 3' SS (Kent and MacMillan, 2002; Donmez et al., 2007). However, in cases where there are multiple 5' or 3' splice sites we do not know their relative positioning. Experiments conducted in the Hertel laboratory show that the spliceosome is committed to using one of the splice sites by A complex and not E complex (Lim and Hertel, 2004; Kotlajich et al., 2009). This is the same point in spliceosome formation at which many inhibitory factors inhibit the exon-definition to intron-definition transition, raising the possibility that the two processes may be linked.

1.8 Total internal reflection fluorescence (TIRF) microscopy

A technique that is capable of answering intriguing problems such as the number of U1 snRNPs that can bind an RNA or the number of U2AF65s that can associate with different pY tracts has recently been developed in the laboratory of Professor Ian Eperon (Cherny et al., 2010). This technique relies upon a prism-based TIRF microscope designed and built within the department by Professor Clive Bagshaw (Conibear and Bagshaw, 2000).

When a beam of light that is passing through a material with a high refractive index comes into contact with a material with a low refractive index total internal reflection can occur. When the angle of incidence is above a critical angle, all the light is reflected; this is total internal reflection. When this occurs, an electromagnetic field, called an evanescent field, is also generated at the interface between the two materials which is the same wavelength as the beam of light that generated it.

(Axelrod et al., 1984)

In a TIRF microscope this evanescent field is used to excite fluorophores that are in very close proximity to the surface of the interface (Figure 1.5). One of the properties of an evanescent field is that it decays exponentially and therefore only fluorophores within a very small region of the sample will be excited (Axelrod et al., 1984), reducing background fluorescence. At very low concentrations of fluorophores this evanescent field-mediated excitation allows investigation of biological processes at the single molecule level.

Fluorescence occurs when a fluorescent molecule releases energy as a photon of light. For this to occur, the fluorescent molecule has to absorb a photon of light, exciting the molecule from a singlet ground state up to an excited state. Once in this state, via internal conversion, the molecule relaxes to the lowest energy level. Once at this level the fluorophore releases the energy in the form of a photon, allowing the fluorophore to return to the ground state. The wavelength of the emission light is longer than that of the excitation light because of a phenomenon known as the Stokes shift. This is where, during the internal conversion phase, energy is lost resulting in the wavelength of light emitted being longer than that of the excitation (Lakowicz, 1999).

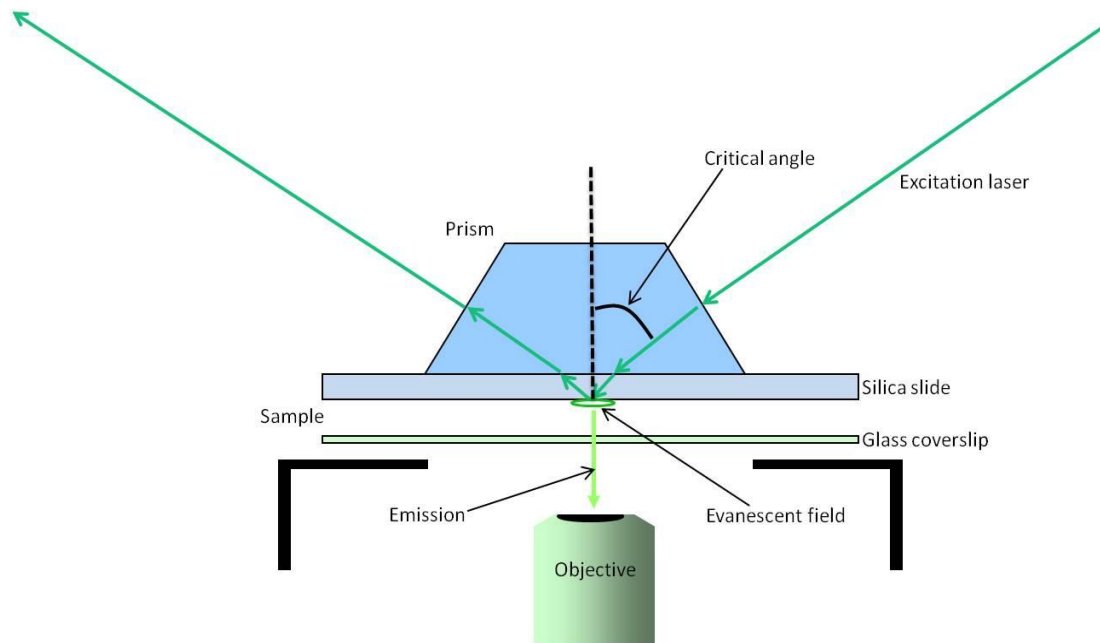


Figure 1.5: Prism-based TIRF microscopy. Illustration showing excitation light hitting the prism and being refracted so that it hits the interface between the silica slide and sample at the critical angle; at or above which, the light is reflected but also leads to the formation of an evanescent field in the sample.

To detect spliceosomal components bound to pre-mRNA fluorescent fusion proteins have been constructed using either the modified enhanced green fluorescent protein (mEGFP) or monomeric cherry (mCherry) fluorophores. Both the mEGFP and mCherry fluorescent protein fluorophores are produced by post-translational autocatalytic cyclization and oxidation reactions (Heim et al., 1994; Cubitt et al., 1995; Niwa et al., 1996; Matz et al., 1999). Although fluorophores can emit photons as described above, this is not the only mechanism by which energy can be released from the fluorophore. Energy can be released through a non-fluorescent pathway in which a triplet state is formed and energy is released as heat. The relative contributions of the fluorescent and non-fluorescent pathways can be described by the quantum yield. The quantum yield is defined as the number of emitted photons relative to the number of absorbed photons (Lakowicz, 1999; Shaner et al., 2005). Formation of triplet states instead of a release of a photon can lead to variations in the intensity that is produced by a fluorophore (Bagshaw and Cherny, 2006). The intensity of a fluorophore can also be affected by other factors such as the polarization effect, where by slow rotation of a fluorophore, relative to the acquisition timeframe, can lead to fluctuations in the intensity level of the fluorophore (Bagshaw and Cherny, 2006).

Another factor that can greatly affect the intensity of a fluorophore is photo-induced isomerisation. This can impact on the intensity of a fluorophore as different isomers of a fluorophore can have different absorption spectra. GFP is known to have at least three ground states which it photoconverts between (Chattoraj et al., 1996; Creemers et al., 1999), and which have different absorbance spectra. These differences in absorbance spectra mean that when a sample is excited using one wavelength of

irradiation you will not detect those fluorophores that are in the other ground states. This leads to a phenomenon known as blinking where by a fluorophore appears to be turning on and off (Bagshaw and Cherny., 2006; Dickson et al., 1997). The different ground states are due to different protonation states and because photon absorbance drives protonation (McAnaney et al., 2005) excitation of a fluorophore can cause alterations in the ground state being occupied. Interestingly, a fluorophore in a dark state can sometimes be induced to fluoresce again by irradiating the fluorophore with a different wavelength light (Dickson et al., 1997; Nifosi et al., 2003) showing the dark state is not permanent.

As well as non-permanent causes of fluorescence loss or variation a fluorophore can also be permanently photobleached. GFP is photobleached because of reactive oxygen species (ROS) that are produced as a bi-product of fluorescence (Greenbaum et al., 2000; Mclean et al., 2009). These ROS are highly reactive and can react with proteins (Davies, 2004) and inactivate the fluorophore. The source of the molecular oxygen to make the ROS is as yet unknown, but there has been speculation that it is trapped in the β Can structure of GFP (Yang et al., 1996; McLean et al., 2009). The lack of interaction between the surrounding environment and the chromophore of GFP (Greenbaum et al., 2000) is thought to be the reason that GFP is relatively resistant to photobleaching (Sullivan and Kay, 1999).

Using TIRF microscopy the number of molecules contained within a fluorescent spot can be investigated. By capturing time-resolved fluorescence for a spot, the number of fluorophores within it can be determined; by looking at the number of times a

photobleaching event occurs (for examples see Funatsu et al., 1995; Leake et al., 2006; Ulbrich and Isacoff, 2007; Shu et al., 2007).

Methods of single molecule analysis similar to those used in this study have recently been used to study: yeast pre-mRNA splicing (Crawford et al., 2007) and spliceosomal assembly (Hoskins et al., 2011). However, we have been the first laboratory to apply this technique to the process of mammalian pre-mRNA splicing. It was used as it enabled the determination of the number of proteins bound to a pre-mRNA without purification of the spliceosomal complexes. This was expected to give us insights into the mechanism of pre-mRNA splicing that no other method could, and is ideal for determining the behaviour of factors that are involved in splice site selection such as U2AF and U1-A.

Chapter 2. Methods

2.1 Procedures related to splicing

- 2.1.1** Preparation of PCR fragments for transcription
- 2.1.2** Radioactive transcription
- 2.1.3** Standard splicing assay
- 2.1.4** Native agarose gel electrophoresis
- 2.1.5** Thin layer chromatography
- 2.1.6** RNase H-mediated degradation of RNA
- 2.1.7** Northern blot analysis of snRNA
- 2.1.8** Denaturing urea polyacrylamide electrophoresis
- 2.1.9** Synthesis of radioactive ladder
- 2.1.10** Silver stain

2.2 Protein analysis techniques

- 2.2.1** Culturing and freezing mammalian cell lines
- 2.2.2** Transfection of HeLa and HEK 293T cells
- 2.2.3** Nuclear extraction
- 2.2.4** Denaturing SDS polyacrylamide electrophoresis
- 2.2.5** Western blot
- 2.2.6** Biotin affinity purification
- 2.2.7** Immunoprecipitation

2.2.8 Ultra-violet cross-linking

2.2.9 RNA extraction

2.2.10 cDNA synthesis

2.2.11 Semi-quantitative PCR

2.3 Methods related to total internal reflection fluorescence microscopy

2.3.1 Cold transcription

2.3.2 Purification of pre-mRNA using ultra violet shadowing

2.3.3 Oligonucleotide labelling of RNA

2.3.4 Sample preparation

2.3.5 Modification of slides

2.3.6 Single molecule imaging of a splicing reaction

2.3.7 Cleaning silica slides

2.3.8 Cleaning glass coverslips

2.3.9 Image analysis

2.3.10 Statistical analysis

2.4 Other techniques

2.4.1 Growing chemically competent bacteria

2.4.2 Transformation of chemically competent bacteria

2.4.3 Mini and maxi preparation of plasmid DNA

2.4.4 Site directed mutagenesis

2.4.5 Agarose gel extraction

2.5 Buffers

2.5.1 Buffers

2.1 Procedures related to splicing

2.1.1 Preparation of PCR fragments for transcription

Full-length versions of the globin constructs were PCR-amplified from plasmids using the primer pairs P16 and P17 or P16 and β GE3no5'ss; versions lacking the 3' end of the RNA were amplified using P16 and Glo C-3'end; the adenovirus (Ad1) constructs were PCR-amplified using Ad1-TrxF and Ad1-TrxR (Table 2.1). The PCR reactions were performed using, GoTaq (Promega), Crimson Taq (Invitrogen), Phusion Hot Start (New England Biolabs) or Red Taq (Sigma) polymerases as per manufacturer's instructions. The PCR fragments were run on agarose gels to ensure specific amplification before being phenol-chlorophorm extracted and ethanol precipitated ready for transcription.

2.1.2 Radioactive transcription

Transcripts labelled with ^{32}P were synthesized using T7 RNA polymerase in 10 μl reactions containing 1x T7 hot transcription buffer, low GTP NTPs (0.5 mM ATP, 0.5 mM CTP, 0.5 mM UTP, 0.05 mM GTP; Promega), 10 mM DTT, 0.22 U diguanosine triphosphate (Roche), 50-150 ng PCR fragment, 5 % v/v RNasin (Promega)/RNasout (Invitrogen), 0.5 μl T7 RNA polymerase and 25-50 μCi [α - ^{32}P] GTP. Reactions were incubated at 37 $^{\circ}\text{C}$ for 1-2 hours, run on a 6 % denaturing polyacrylamide gel and extracted. RNAs were eluted from the gel slice in RNA elution buffer before being ethanol precipitated and resuspended ready for use.

Labelling of the transcripts with [α - ^{32}P] UTP was achieved by replacing the [α - ^{32}P] GTP with [α - ^{32}P] UTP and the low GTP rNTPs with low UTP NTPs (0.5 mM ATP, 0.5 mM CTP, 0.05 mM UTP, 0.5 mM GTP).

<u>Name</u>	<u>Sequence</u>
P16	AAATTAATACGACTCACTATAGGGCTGCTGGTTGTCTACCCA
P17	ACTTACCTGCCAAAATGATGAGACAGCACAAT
βG E3 no 5'ss	CCAAAATGATGAGACAGCACAAT
Glo C -3'end	GTTAGCAGAGTCGACCATAT
Ad1 Trx F	TAATACGACTCACTATAGGGAACAAAAGCTTGTATGCCT
Ad1 Trx R	GATCCAAGAGTACTGGAAAGACCGCGA
α3EDNA	CTCTGGGCCTAACCAGCA
SF2-F	TGGGGGTGTTTGTGGTTATC
SF2-R	CCAGTGAGCCCTCTCCAATC
α-U1	GCCAGGUAAGUAU -BIOTIN
α-U6	CUGUGUAUCGUUCCAAUUUU

Table 2.1: Sequences of Oligonucleotides. Bold lettering refers to 2'O methyl modified nucleotides; BIOTIN refers to biotin group.

2.1.3 Standard splicing assay

Splicing reactions containing 1.5 mM ATP, 3.2 mM MgCl₂, 20 mM phosphocreatine, 50 mM potassium glutamate, radiolabeled pre-mRNA and 50 % NE (Cilbiotech, or as described in Section 2.2.3, referred to as 'home-made') were assembled and incubated at 30 °C for the period indicated. Time points were taken and stored on dry ice or at -80 °C until the time course was completed. Samples were treated with proteinase K (0.4 mg/ml) in PK buffer for 10-30 minutes before being ethanol precipitated and run on a denaturing polyacrylamide gel.

2.1.4 Native agarose gel electrophoresis

Native agarose gel electrophoresis was carried out as in Das and Reed, (1999). To resolve E complex from H complex, NE was pre-incubated at 30 °C for 30 minutes in order to deplete the endogenous ATP. Standard splicing reactions lacking ATP and CrPi were then assembled using the pre-depleted NE, and incubated for the time indicated. Early time-points were sometimes stored at -80 °C or in liquid N₂ until the time course was completed. An equal volume of TG loading buffer was added to each sample prior to loading on a 1.5 % LMP agarose gel, which was run in 50 mM Tris, 50 mM Glycine for up to four hours at 4 °C. The LMP agarose was supplied by Invitrogen.

To deplete ATP using Hexokinase and glucose 250 U/ml Hexokinase and 10 mg/ml glucose was added to 10 µl of NE. This reaction was then incubated at 30 °C for 30 minutes before being used as part of a splicing reaction.

To inhibit the activities of phosphatases during ATP depletion Phos-stop phosphatase inhibitor (Roche) was added to nuclear extract at a final concentration of 2 X before incubation at 30 °C for 30 minutes.

To resolve the A, B and C complexes from H complex, standard splicing reactions were assembled and incubated for the indicated time. Time-points were incubated at room temperature for 30 minutes in the presence of 0.8 mg/ml heparin. An equal volume of TG loading buffer was added to each sample and they were run on a 2 % LMP agarose gel at 4 °C in TG running buffer. Two exon constructs were electrophoresed for three and a half hours, three exon constructs for seven hours.

The spliceosome was stalled at A complex by pre-incubating the NE in the presence of 0.3-0.6 mM anacardic acid (Calbiochem), 0.4-1 mM garcinol (Enzo Life Sciences) or 1 μ M anti U6 oligonucleotide (Table 2.1; Eurogentec) at 30 $^{\circ}$ C for 15 minutes before adding it to a standard splicing reaction (Kuhn et al., 2009, Donmez et al., 2007).

2.1.5 Thin layer chromatography

To analyse the depletion of ATP from NE approximately 2.5 μ Ci [γ - 32 P] ATP was added to each 10 μ l sample of NE. After incubation at 30 $^{\circ}$ C, 5 μ l timepoints were added to 2.5 μ l 4 M potassium acetate, 10 μ l 7 % perchloric Acid and 5 μ l water to increase the pH to between 3.5 and 4. The samples were then loaded onto a poly-ethyleneimine-cellulose thin layer chromatography plate (Sigma) and chromatographed in 4 M potassium acetate until the front had covered 80 % of the plate. The plate was then dried and exposed to a phosphor screen.

2.1.6 RNaseH-mediated degradation of RNA

To degrade the snRNA in nuclear extract, standard 40 μ l splicing reactions were assembled, with the addition of 0.05 % NP40 but without the pre-mRNA. The snRNA-specific oligonucleotides and 4 U of RNase H (New England Biolabs) were added to the reaction and incubated at 30 $^{\circ}$ C for 45 minutes to digest the snRNA. Splicing was then carried out using this snRNA depleted extract.

To analyse products formed in a splicing reaction the RNA was first purified and resuspended in 5 μ l water. One-fifth of this was then added to a mixture containing 100 mM NaCl, 10 mM Hepes pH7.5 and 10 pmol/ μ l oligonucleotide and made to 5 μ l with water. This mixture was incubated at 80 $^{\circ}$ C for 5 minutes, room temperature for

5 minutes and then placed on ice for 5 minutes to allow annealing of the oligonucleotide to the RNA. The oligonucleotide-annealed RNA was added to 4 U RNaseH in 1x RNase H buffer, made up to 10 µl with water and then incubated at 30 °C to allow cleavage. Formamide dyes were added and the reactions were run on a 12 % denaturing urea polyacrylamide gel.

To cut the CEC RNA in the central exon, splicing reactions were carried out in the presence of 1.5 U/10 µl RNase H. Reactions were incubated at 30 °C, time points were taken and stored in liquid N₂. Samples were transferred to room temperature and α3E-DNA oligonucleotide (Table 2.1) was immediately added to a concentration of 17 pmol/µl; they were then incubated at 30 °C for 5 minutes. Either denaturing polyacrylamide or native agarose gel electrophoresis were used for analysis.

2.1.7 Northern blot analysis of snRNA

Total RNA was isolated from NE by proteinase K-treatment and ethanol precipitation before resuspension in formamide dyes. Samples were loaded onto 6 % denaturing, urea containing polyacrylamide gels and run until the bromophenol blue was at the bottom of the gel. The gel was then soaked in 5X SSC, pH 7.4, for 10 minutes before transfer onto an NX membrane (Amersham) overnight. The membrane was then baked for 2 hours at 80 °C and then incubated at 37 °C in 1X Church and Gilbert solution for 2 hours. Riboprobes against the snRNAs (Blencowe et al., 1989) were transcribed as in Section 2.1.2 and added to the mixture and incubated rotating at 37 °C overnight. The membrane was then washed three times in 1X SSC, pH 7.4, for 5 minutes at room temperature. This was followed by one wash in 1X SSC, pH 7.4, 0.1 %

SDS at 65 °C for 30 minutes. The membrane was then exposed to a Phosphor screen to visualise the RNAs.

This protocol was adapted for native agarose gels by washing the agarose gel after electrophoresis with 0.25 M HCl for 10 minutes to remove the protein, before washing the gel in 5X SSC, pH7.4, for 10 minutes.

2.1.8 Denaturing urea polyacrylamide gel electrophoresis

Gels ranged in volume from 5 ml to 80 ml and contained 7 M Urea, 4-12 % acrylamide 19:1 bisacrylamide solution (National Diagnostics) and 1X TBE. Polymerisation was catalysed by the addition of ammonium persulfate and TEMED (Sigma). Samples in formamide dyes were loaded onto the gel and run for various times depending on the RNA used. Gels were dried and exposed to Phosphor screens for analysis.

2.1.9 Synthesis of radioactive ladder

Radioactive DNA ladders were made by digesting the plasmid pBR322 with the restriction enzymes HpaII, HaeIII or AluI at 37 °C for 30 minutes. Markers made using plasmid digested with HpaII will be referred to as M₁, HaeIII will be M₂ and AluI will be M₃. The sizes of the bands produced by each digestion are given in Table 2.2. After digestion the DNAs were treated with antarctic phosphatase at 37 °C for 15 minutes followed by inactivation at 65 °C for 15 minutes. One-tenth of the total digest was added to a T4 polynucleotide kinase reaction containing [γ -³²P] ATP and incubated for 30 minutes at 37 °C. Free nucleotides were removed by centrifugation through a G50 spin column (GE Healthcare), the mixture was diluted in 5 volumes of formamide dyes,

incubated at 80 °C for 2-5 minutes and loaded on a denaturing urea polyacrylamide gel.

<u>Enzyme</u>	<u>Band sizes (base pairs)</u>
Hpa II (M ₁)	622, 527, 404, 307, 242, 238, 217, 201, 190, 180, 160, 147, 123, 110, 90, 76, 67.
Hae III (M ₂)	587, 540, 504, 458, 434, 267, 234, 213, 192, 184, 124, 123, 104, 89, 80, 64, 57, 51, 21, 18, 11, 7.
Alu I (M ₃)	908, 659, 656, 521, 403, 281, 257, 226, 100

Table 2.2: Band sizes of pBR322 digests. M₁, M₂ and M₃ refer to the names given to each ladder later in text.

2.1.10 Silver stain

Total RNA from NE were first run on a denaturing urea polyacrylamide gel. The gel was treated in the following manner: it was soaked in 50 % acetone, 0.37 % formaldehyde, 0.04 % trichloroacetic acid for 5 minutes; washed in water three times for 5 seconds, then 5 minutes; soaked in 50 % acetone for 5 minutes; washed twice in water for 5 seconds; incubated in 0.01 % Na₂S₃O₂ for 1 minute; washed twice with water for 5 seconds; incubated in 16.5 mM AgNO₃ and 0.74 % formaldehyde for 8 minutes; washed twice in water. The gel was then exposed for about 10 seconds in 172 mM Na₂CO₃, 0.27 % Na₂S₃O₂, 0.015 % formaldehyde; exposure was stopped using 1 % acetic acid.

2.2 Protein analysis techniques

2.2.1 Culturing and freezing cell lines

Tissue culture cells, HeLa and HEK 293T were maintained at 37 °C with 5 % CO₂, in 1x DMEM containing 1x penicillin/streptomycin solution and 10 % FBS. Cells were split regularly and kept until passage number 30. Cell stocks were frozen at 2x10⁶ cells/ml in 70 % DMEM, 20 % FBS and 10 % sterile DMSO. Stocks were frozen slowly in three stages by placing at -20 °C for 2 hours then -80 °C overnight, and finally transferred to liquid nitrogen for long term storage. All solutions were purchased from Invitrogen.

2.2.2 Transfection of HeLa and HEK 293T cells

For transfection with Fugene 6 (Roche), HeLa cells were seeded at 4x10⁶ cells/176 cm² dish or 5x10⁵ cells per well of a six-well plate, whilst HEK 293T cells were plated at 6x10⁶ cells/ 176 cm² dish and cultured for 24 hours prior to transfection according to the manufacturer's instructions. Cells were then grown for a further 48 hours for HeLa cells or 24 hours for HEK 293T cells before harvesting.

For transfection using the calcium chloride procedure, HEK 293T cells were plated at 6x10⁶ cells/ 176cm² dish and cultured for 24 hours. The media was replaced 4 hours prior to transfection. Plasmid DNA (2-16µg) was added to 1.67 ml of 1x CaCl₂ solution and then mixed with 1.67 ml of ice cold 1x HBS and incubated on ice for ten minutes. This mixture was then carefully pipetted onto the cells and cultured for 24 hours. The cells were shocked for three minutes at room temperature in a freshly made 25 % DMSO, 75 % DMEM solution pre-warmed to 37 °C and then washed three times at

room temperature in DMEM media. They were cultured for a further 24 hours and harvested.

2.2.3 Nuclear extraction

Nuclear extracts were prepared as in Lee et al. (1988). Cells were harvested from two 176 cm² dishes using a cell scraper and spun at 150 g for eight minutes. Cell pellets were resuspended in 1 packed cell volume of buffer A and allowed to swell on ice for 15 minutes. Cells were then lysed either by vortexing for 10 seconds (Vortex-genie 2) or by forcing the cells through a 25g needle six times. The homogenate was centrifuged for 1 minute at 9250 g to collect the nuclei. The nuclear pellet was resuspended in two-thirds of the original packed cell volume of buffer C and then incubated at 4 °C on a magnetic stirrer with a micro-flea for 30 minutes. The nuclear debris was precipitated by centrifugation at 9250 g for 5 minutes at 4 °C. The supernatant was dialyzed against buffer D at 4 °C by placing 50 µl aliquots on floating 0.025 µm filters (Millipore). The nuclear extract was collected from the filters, snap-frozen in liquid nitrogen and stored at -80 °C.

2.2.4 Denaturing SDS polyacrylamide electrophoresis

A polyacrylamide stacking gel (4 % [37.5 acrylamide : 1 bisacrylamide solution; National Diagnostics], 0.125 M Tris-HCl pH 6.8, 0.1 % SDS, 0.1 % ammonium persulfate, 0.13 % Temed) was laid over a 10 or 12 % polyacrylamide resolving gel (10 or 12 % [37.5 acrylamide : 1 bisacrylamide solution], 0.37 M Tris-HCl pH 8.6, 0.1 % SDS, 0.1 % ammonium persulfate, 0.13 % Temed) using the mini-Protean kit (Biorad). Samples were loaded onto the gel in 1x SDS sample buffer and run through the stacking gel at

80 V; the voltage was then increased to 150 V to run the samples through the resolving gel. Gels were run in 1x SDS-PAGE buffer.

2.2.5 Western blot

Gels containing separated protein samples were transferred in 1x transfer buffer using a semi-dry blotting apparatus (Biometra or Biorad) onto nitrocellulose membrane.

Once transferred, the membranes were incubated in blocking buffer overnight at 4 °C.

Membranes were then incubated with the required primary antibody (Table 2.3)

diluted 1:1000 in blocking buffer for 1 hour at room temperature. Membranes were

washed three times in blocking buffer for 10 minutes before being incubated with

protein A/G conjugated to horseradish peroxidase (A/G HRP; Table 2.3) which had

been diluted between 1:1000 and 1:5000 in blocking buffer. Membranes were

washed again three times in blocking buffer before application of Western Dura Super-

signal horseradish peroxidase reagent (Pierce). Blots were covered in cling-film and

exposed to X-ray film for between 2 seconds and 5 minutes.

For fluorescent Western blots, the blocking buffer was replaced with the Odyssey

blocking buffer (LiCor) diluted 1:1 in TBS. The A/G HRP was replaced by fluorescent

secondary α -rabbit or α -mouse antibodies (Table 2.3), diluted to 1:15000 in Odyssey

blocking buffer. Fluorescent blots were imaged and quantified using the Odyssey

scanner (LiCor).

<u>Antibody</u>	<u>Name</u>	<u>Source</u>	<u>Supplier</u>
α U1A (a)	-	Mouse (monoclonal)	I.Mattaj, Kambach and Mattaj, (1992)
α U1A (b)	Ab55751	Mouse (monoclonal)	AbCam (Cambridge, UK)
α U2AF35	Ab1022	Rabbit	Dundee Cell Products (Dundee, UK)
α U2AF65	mAB MC3	Mouse (monoclonal)	M. Carmo Fonseca, Gama-Carvalho et al., (1997)
α EGFP	Living Colours Full Length Av Polyclonal Ab	Rabbit (polyclonal)	Clontech (St-Germain-en-Laye, France)
α DSRed	Living Colours DSRed Polyclonal Ab	Rabbit (polyclonal)	Clontech (St-Germain-en-Laye, France)
α Rabbit	IRDye 680CW α -Rabbit	Goat (monoclonal)	Li-Cor (Lincoln, Neb, USA)
α Mouse	IRDye 800CW α -Mouse	Goat (monoclonal)	Li-Cor (Lincoln, Neb, USA)
Protein A/G HRP	32490	Protein A/G HRP conjugated	Thermo Fisher Scientific (Rockford, IL, USA)

Table 2.3: Primary and Secondary Antibodies. Antibodies used for Western blot, Fluorescent Western blot and Immunoprecipitation.

2.2.6 Biotin affinity purification

To precipitate components of U1 snRNPs, 15 μ l standard splicing reactions were set up. Prior to incubation at 30 $^{\circ}$ C for 20 minutes, the α -U1 snRNA oligonucleotide (Table 2.1) was added to a final concentration of 200 nM. Any precipitate was removed by centrifugation at 14000 r.p.m. for 1 minute. The supernatant was added to 10 μ l neutravidin beads (Pierce) that had been pre-washed with SDS wash buffer to remove loose neutravidin and made up to 50 μ l using Buffer D. The mixture was incubated at 4 $^{\circ}$ C rotating for 45 minutes to allow the biotin to bind. Beads were washed three

times with 1 ml of FSP buffer containing 100 mM NaCl, followed by three washes in 1 ml of FSP buffer. The beads were then boiled for 5 minutes in 1x SDS sample buffer and the protein-containing supernatant was then loaded onto a 12 % SDS-polyacrylamide gel.

To precipitate components bound to biotinylated RNAs, the protocol was modified as follows: instead of adding α -U1 snRNA oligonucleotide to a 15 μ l reaction, 50 nM RNA was added and incubated in 100 μ l standard splicing reaction; 20 μ l of neutravidin beads were used instead of 10 μ l.

2.2.7 Immunoprecipitation

To immunoprecipitate RNAs bound to protein, 50 μ l standard splicing reactions were incubated at 30 $^{\circ}$ C to allow complex formation. To pre-clear the reactions they were then added to 50 μ l protein G bead slurry (Pierce) in 300 μ l IP200 buffer and incubated for two hours, rotating at 4 $^{\circ}$ C. Whilst the samples were being incubated, 2.1 μ g α -EGFP or α -DSRed antibody (Table 2.3) was bound to 75 μ l of protein G slurry in 1 ml of PBS supplemented with 0.1 mg/ml BSA by incubation for two hours, rotating at 4 $^{\circ}$ C. The antibody bound beads were washed twice in 1 ml of PBS and once in 1 ml IP200. Pre-cleared splicing reactions were then added to 15 μ l antibody-bound beads and incubated, rotating at 4 $^{\circ}$ C for two hours. These were then washed three times in 1 ml of IP200 to remove non-specifically bound protein-RNA complexes. The RNA was purified by proteinase K treatment and ethanol precipitation, resuspended in formamide dyes and run on a 6 % denaturing urea polyacrylamide gel.

2.2.8 Ultra-violet cross-linking

Ultra-violet cross-linking of mEGFP-U2AF65 to the rat α -tropomyosin RNA fragment was done in a 10 μ l standard splicing reaction where 250 mM KCl was substituted for the 50 mM potassium glutamate. Splicing reactions were incubated at 30 $^{\circ}$ C for 30 minutes to allow complex formation before being cross-linked using a UVP Spotcure set to 254 nm for one minute. Each reaction was incubated with 1 μ l of RNase A for 30 minutes at 37 $^{\circ}$ C. Samples were added to 10 μ l 2x SDS sample buffer and run on a 10 % SDS polyacrylamide gel. Gels were then transferred onto nitrocellulose membrane using the Biometra semi-dry blotting system and imaged using a Phosphor screen.

2.2.9 RNA extraction

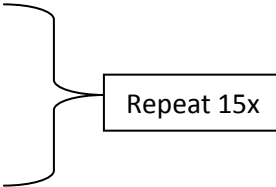
Cells were stripped from the surface of a single well from a 6-well plate using 500 μ l Tri-reagent (Sigma), transferred to a 1.5 ml Eppendorf tube and incubated at room temperature for 5 minutes. The sample was shaken vigorously after the addition of 100 μ l chloroform, left at room temperature for 15 minutes and then centrifuged at 6870 g for 15 minutes at 4 $^{\circ}$ C. The aqueous phase was added to 250 μ l of isopropanol, mixed, left at room temperature for 10 minutes and then precipitated by centrifugation at 6870 g for 10 minutes at 4 $^{\circ}$ C. The pellet was washed in 100 % ethanol before being dried and resuspended in 40 μ l water. The RNA was treated with 0.1 U/ μ l of DNase (Promega) for 20 minutes before being phenol-chlorophorm extracted and ethanol precipitated. The purified RNA was then used for cRNA synthesis.

2.2.10 cDNA synthesis

Purified RNA (2 µg) was mixed with 1 µg Random Primers (Promega) and made up to 15 µl with water. This was heated to 70 °C for 15 minutes before quick-cooling on ice for 10 minutes. The reaction mixture was divided into two, with half being treated with reverse transcriptase and the rest used as a control. Reactions contained either 1 µl MMLV-RT (RNase H-) or 1 µl water and 1x MMLV-RT buffer, 0.5 mM dNTPs, 0.5 µl RNasin (Promega). The reactions were incubated at 37 °C for 1 hour to synthesize the cDNA prior to freezing.

2.2.11 Semi-quantitative PCR

The PCR reactions were performed using GoTaq polymerase (Promega) and the SF2-F and SF2-R primers (Table 2.1) using the following cycle in a thermocycler.

- 94 °C 2 minutes
 - 94 °C 30 seconds
 - 55 °C 30 seconds
 - 72 °C 1 minute
 - 72 °C 5 minutes
- 
- Repeat 15x

Samples were then run on 1.5 % agarose gels and post-stained with ethidium bromide for 15 minutes. The gels were then imaged using a Typhoon Imager and analysed using Optiquant software.

2.3 Methods related to total internal reflection fluorescence microscopy

2.3.1 Cold transcription

Reactions containing 1x cold transcription buffer, 10 mM DTT, 4 mM NTPs (Promega), 1 µg transcription template, 5 % v/v T7 RNA polymerase and 40-200 U RNasout (invitrogen), were incubated at 37 °C for 4 hours. They were then DNase-treated as described in Section 2.2.9, passed through a S-300 column (GE Healthcare), phenol-chlorophorm extracted and ethanol precipitated before being dried and resuspended in 25 µl of water. Concentrations were calculated based on the absorbance at 260 nm. The protocol was modified to make biotinylated RNA by addition of 1 µM biotin-16-UTP (Roche) to the reaction.

2.3.2 Purification of pre-mRNA using ultra violet shadowing

The transcription was carried out as in section 2.3.1 and then added to Formamide dyes before being run on a denaturing urea 6 % polyacrylamide gel. Once run the gel was wrapped in cling film and placed on a Silica gel 60 T₂₅₄ TLC plate (Merck). The gel was then exposed to 254 nm ultra-violet light, the RNA absorbed this light and as such the position of the RNA within the gel could be seen with the naked eye. The RNA was then cut out of the gel and eluted in elution buffer overnight. The RNA was then ethanol precipitated, dried and resuspended ready for use.

2.3.3 Oligonucleotide labelling of cold RNA

The transcript was annealed to the fluorescently-labelled oligonucleotide (Eurogentec) by mixing 1 µM of fluorescent oligonucleotide with a series of RNA concentrations,

typically ranging from 0.85 μM to 1.5 μM , in annealing buffer and heating to 80 $^{\circ}\text{C}$ for 5 minutes followed by slow cooling to 40 $^{\circ}\text{C}$ in either a PCR machine or hot block. Oligonucleotide-annealed RNAs were cooled further on ice for 1-2 hours. The samples were run on a 6 % polyacrylamide gel in OL gel dyes; the amount of free oligonucleotide in the annealing reaction was assessed by detection of the fluorophore using a Typhoon imager. A sample that contained minimal free oligonucleotide was then used for experiments.

2.3.4 Sample preparation for TIRF microscopy

Splicing reactions contained 50 nM oligonucleotide-labelled RNA, 45 % NE, 3 mM MgCl_2 , 1.36 mM rATP, 18 mM CrPi, 1x D-glu, 36 mM Hepes-KOH pH 7.5, 1 % v/v RNasin or RNasout.

For reactions carried out in the absence of ATP, ATP and phosphocreatine were omitted from the reaction and the NE was pre-incubated at 30 $^{\circ}\text{C}$ for 30 minutes to deplete ATP. Reactions to be inhibited with anacardic acid or α -U6 oligonucleotide were pre-incubated in the absence of RNA with 0.3 mM or 0.6 mM anacardic acid or 1 μM α -U6 oligonucleotide respectively.

Complexes were assembled by incubation at 30 $^{\circ}\text{C}$ for 30 minutes. They were then diluted in dilution buffer until the RNA concentration was approximately 3 pM. The sample was then ready to be loaded into the flow cell for imaging.

2.3.5 Modification of Slides for TIRF microscopy

Flow cells were assembled essentially as in Conibear and Bagshaw (2000). Borosilicate glass coverslips (number 1), approximately 140 μm thick, were attached to brass

mounts using double-sided sticky tape. Silica slides were then attached onto the coverslips, again using double-sided sticky tape, to create a flow cell between the coverslip and the slide. Figure 2.1 shows a diagrammatic representation of the flow cell.

The surface of the flow cell was then modified so that it will bind RNA sufficiently (Cherny et al., 2010). First, 30 μ l of biotinylated BSA solution (10 μ g/ml in dilution buffer) was injected into the flow cell and incubated at room temperature for 10-15 minutes; any biotinylated BSA that was not absorbed onto the surface was then washed out using 30 μ l of dilution buffer. To create a binding site for the biotin on the oligonucleotide annealed RNA, 30 μ l of streptavidin (10 μ g/ml in dilution buffer) was then injected and the flow cell was incubated at room temperature for 10-15 minutes. Free streptavidin was then washed out using dilution buffer.

2.3.6 Single molecule imaging of a splicing reaction

Single molecule images were obtained using a custom-built prism-based TIRF microscope (Conibear and Bagshaw, 2000) using conditions very similar to those used by Cherny et al. (2010). The diluted samples were loaded into the flow cell and when enough fluorescent spots could be seen, the sample was flushed with imaging buffer containing the oxygen scavenging system, 50 nM protocatechuate dioxygenase (Sigma) and 2.5 mM protocatechuic Acid (Aitken et al., 2008) to increase the life-time of the Cy5 and Cy3 fluorophores.

For each experiment appropriate lasers, dichroic mirrors and emission filters were selected based on the fluorophores present in the sample (Table 2.4 and 2.5).

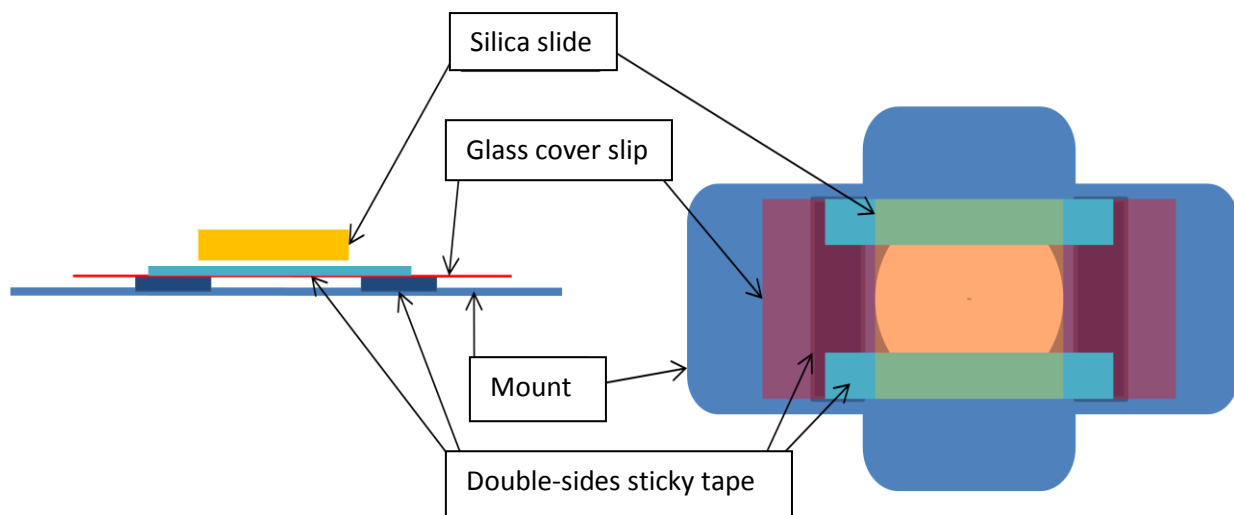


Figure 2.1: The arrangement of the slide for single molecule experiments

Upon excitation with an appropriate laser, emission fluorescence from the fluorophores was collected using a 63x 1.2 N.A. Zeiss C-apochromat water-immersion lens. The emission was then sent through an aperture to halve the size of the image and the beam was then split by an appropriate dichroic mirror; any excitation light was filtered out using emission filters for the appropriate wavelength. The emission from both fluorophores was detected on a single chip because the image has been halved in size. The camera used for detection was an Ixon DU887 emCCD camera from Andor Technologies UK. The detector chip was setup to acquire an image every 200ms to allow time-resolved fluorescence measurements. For experiments with Cy5 labelled RNA and mEGFP fusion proteins, 20-30 frames of emission from the Cy5 fluorophore were captured, followed by 275-375 frames of emission from the mEGFP fluorophore. Experiments with the Cy5-Cy3-labelled RNAs were imaged for 50 frames under each emission condition, whilst in experiments with the mEGFP and mCherry fusion proteins 100 frames of emission for each were captured.

<u>Fluorophore</u>	<u>Laser Type</u>	<u>Excitation Wavelength (nm)</u>	<u>Power at prism (W/cm²)</u>	<u>Emission filter (Omega)</u>
mEGFP	Argon-Ion	488	~100	510DF23
mCherry	Nd:YAG (1)	561	~80	595AF6D
Cy5	He-Ne (2)	633	~50	670DF40
Cy3	DPSS (3)	532	~100	580DF30

Table 2.4: Laser and emission filter requirements for fluorophores. Abbreviation: 1: neodymium-doped yttrium aluminium garnet, 2: Helium-Neon, 3: Diode-pumped solid-state

<u>Fluorophore emission wavelengths to be split</u>	<u>Dichroic Mirror (Omega)</u>
mEGFP and Cy5	540DCLP
mEGFP and mCherry	540DCLP
Cy5 and Cy3	610DRLP

Table 2.5: Dichroic mirrors used to separate emission wavelengths

2.3.7 Cleaning silica slides

Silica slides were re-used and therefore had to be cleaned thoroughly to remove all of the fluorophores from the surface. Slides were soaked in 10 % Alconox detergent and sonicated for 15 minutes in a sonicating water bath; sonication was repeated 4-5x with intervals of at least 30 minutes to allow cooling. Detergent was removed by rinsing in water, followed by sonication in water 5x as before. The water was then replaced with 1 M KOH and the sonicated 5x as before. The 1 M KOH was removed and the slides were rinsed several times with high purity (MilliQ) water and then sonicated 5x as before. Slides were stored in high quality water until use.

2.3.8 Cleaning glass coverslips

The borosilicate glass coverslips were used only once but to ensure that no contamination occurred, they were soaked in 1 M KOH for 1-2 hours and then rinsed in high quality running water. Traces of water were removed by vacuum aspiration; slides were left to air dry overnight before use.

2.3.9 Image analysis for mEGFP-U1-A co-localisation data

For each set of time-resolved images, an accumulated image of Cy5 and mEGFP signals was acquired by using Image J to collect a 'sum image' of approximately 20 frames. Pseudo-colour was then added to the image so that Cy5 signal was red and mEGFP signal was green. The two images were then overlaid to reveal co-localisation. Overlaying was based on a sample of control fluorescent beads, which fluoresce in both channels (Invitrogen, Trans FluoSphere (488/605)). Bleaching profiles were then acquired from the original raw data for all co-localised spots in each frame, using

Image J. An area of approximately 8x8 pixels was selected for each spot in order to limit the level of background interference. Co-localised spots, where the mEGFP spots appeared part of the way through the excitation with the 488 nm laser, were not analysed. Bleaching curves were then saved in an Excel spread sheet and the traces were classified as 1 to 3-mers for all experiments. In each experiment there were spots that could not be classified: these could have been due to fluctuations in intensity obscuring the true number of fluorophores or a lack of bleaching. As the proportion of spots having three-step bleaching curves was very low they were included in the unclassified results for the statistical analysis.

2.3.10 Statistical analysis

For analysis of the mEGFP-U1-A data, fluorescent western blots were used to calculate the ratio between endogenous and fluorescently-labelled U1-A. This ratio was then used to calculate a binomial distribution in order to ascertain the expected values for two U1 snRNPs being bound per RNA. The assumption was made that firstly the presence of the mEGFP tag did not hinder the ability of U1-A to bind U1 snRNA and secondly that its presence within the U1 snRNP did not hinder binding to RNA. These assumptions were made because U1-A binds U1 snRNA in a region far away from to its 5' end (Weber et al., 2010) and has been shown to unnecessary for RNA binding (Heinrichs et al., 1990). The expected values were then compared to the observed values using a Chi-Square test .

2.4 Other techniques

2.4.1 Growing chemically competent cells

Bacterial cultures (5 ml) were grown overnight without antibiotics. The culture was then diluted 1:20 in 100 ml LB and grown to an OD of 0.5 before harvesting by centrifugation at 4850 g for 5 minutes at 4 °C. The pellet was resuspended in 25 ml of 0.1 M CaCl₂ and placed on ice for 30 minutes. The cells were collected by centrifugation at 4850 g for 5 minutes at 4 °C before resuspension in 4 ml of 0.1 M CaCl₂, 15 % glycerol. Cells were then snap-frozen in a solution of dry-ice and industrial methylated spirits in 25-50 µl aliquots.

2.4.2 Transformation of chemically competent cells

Plasmid DNA (10 ng-50 ng) was added to 25-50 µl of chemically competent cells (TOP-10 or XL-1 blue), and incubated on ice for 5 minutes. Cells were then heat-shocked for 30 seconds at 42 °C, transferred to ice and 250 µl of SOC (Invitrogen) or LB media was added before incubation, with shaking at 215 r.p.m., at 37 °C for 1 hour. The cells were then spread onto LB-agar plates containing appropriate antibiotics and grown overnight at 37 °C.

2.4.3 Mini and Maxi prep

Plasmid mini-preps were carried out either using a kit (Qiagen) or as follows. A single colony of the transformed *E. coli* was grown overnight in a 5 ml culture before centrifugation at 4850 g for 5 minutes. The pellet was resuspended in 200 µl of Solution I and incubated at room temperature for 10 minutes. Freshly made Solution II (400 µl) was then added, mixed and incubated on ice for 5 minutes; this was

followed by the addition of 300 µl of Solution III, mixing and incubation on ice for 10 minutes. Samples were then spun at 10730 g for 10 minutes to precipitate cellular debris and 400 µl of the supernatant was added to 1 ml 100 % ethanol, mixed by hand, and spun at 10730 g for 10 minutes. The DNA pellet was washed once in 100 % ethanol and spun again as before. The pellet was then dried for about 10 minutes, resuspended and extracted with phenol-chloroform using standard procedures.

The purification of plasmid DNA for transfection of mammalian cell lines was carried out using the endotoxin-free maxi-prep kit made by Qiagen.

2.4.4 Site Directed Mutagenesis

Many of the mutant constructs were made using a standardised site directed mutagenesis protocol. Complementary pairs of primers were designed to prime in opposite directions such that there were 20 complementary nucleotides either side of the mismatch. Amplification by PCR was carried out using the Pfu-Turbo polymerase (Stratagene), 2.5 µM oligonucleotide and 10-100 ng plasmid DNA. Samples were digested with DpnI to remove parental plasmid. The reactions were transformed into TOP-10 cells and plated onto appropriate antibiotic-containing LB-agar plates.

2.4.5 Agarose Gel Extraction

DNA fragments were extracted from agarose gel slices either using a gel extraction kit (Qiagen) as per the manufacturer's instructions or using NaI. To extract fragments using NaI, gel pieces were added to 2.5 volumes of 6.6 M NaI assuming 1 mg of gel = 1 µl 6.6M NaI. Gel slices were then incubated at 50 °C to melt the gel fragment. 3 volumes of room temperature GB binding buffer (Abgene) were then added and mixed

well. The mixture was then added to an Abgene spin filter, spun through the column and then washed with GW wash solution (Abgene) to remove residual NaI. 30 µl TE.1 was then added to the column and the DNA was removed by centrifugation. The DNA was then phenol-chlorophorm extracted and ethanol-precipitated before being used.

2.5 Buffers

- Annealing buffer: 10 mM Hepes pH 8.0, 100 mM NaCl
- Blocking buffer: 25 mM Tris base, 150 mM NaCl, 2 mM KCl, 5 % w/v dried milk powder (Marvel), 0.01 % v/v Tween
- Buffer A: 10 mM Hepes pH 8.0, 1.5 mM MgCl₂, 10 mM KCl, 1 mM DTT
- Buffer C: 20 mM Hepes pH 8.0, 25 % v/v glycerol, 420 mM NaCl, 0.2 mM EDTA, 1 mM DTT
- Buffer D: 20 mM Hepes pH 8.0, 10 % v/v glycerol, 100 mM KCl, 0.2 mM EDTA, 1 mM DTT
- 1X CaCl₂ solution: 860 µM Tris-HCl pH 7.5, 86 µM EDTA, 260 mM CaCl₂
- Church and Gilbert solution: 22.55 g Na₂HPO₄, 10.92 g NaH₂PO₄, 35 g SDS in 500ml H₂O
- Dilution buffer: 10 mM Hepes pH 7.5, 100 mM NaCl, 0.2 v/v RNasout
- 5X D glu: 177.8 mg potassium glutamate, 4.8 µl 0.5 M EDTA, 600 µl glycerol, 5.4 ml water
- Formamide dyes: 90 % v/v formamide, 50 mM EDTA, bromophenol blue, xylene cyanol
- FSP: 20 mM Tris-HCl pH 7.5, 60 mM KCl, 2.5 mM EDTA, 0.1 % v/v Triton X100
- 1X HBS: 135 mM NaCl, 5 mM KCl, 0.7 mM mM Na₂HPO₄, 5.5 mM glucose, 21 mM Hepes pH 7.05
- Imaging buffer: 50 nM protocatachuate dioxygenase (Sigma), 2.5 mM protocatachuic acid (Spectrum chemicals), 100 mM NaCl,

10 mM Hepes pH 7.5, 0.5 % v/v RNasout

- IP200: 20 mM Hepes pH8.0, 1.5 mM $MgCl_2$, 0.05 % NP40, 200 mM NaCl
- LB media: 10 g Bacto-tryptone, 5 g Yeast extract, 10 g NaCl in 1 L H_2O
- OL dyes: 90 % v/v Formamide, 50 mM EDTA
- PBS: 136.89 mM NaCl, 2.68 mM KCl, 10.14 mM Na_2HPO_4 , 1.76 mM KH_2PO_4
- PK buffer: 100 mM Tris-HCl pH 7.5, 12.5 mM EDTA, 150 mM NaCl, 1 % w/v SDS
- RNA elution buffer: 1 mM EDTA, 0.2 % w/v SDS, 0.5 M sodium acetate pH 4.0
- RNase H buffer: 50 mM Tris-HCl pH 8, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT
- 1X SDS PAGE running buffer: 25 mM Tris base, 192 mM glycine, 0.02 % w/v SDS
- 1X SDS PAGE sample buffer: 250 mM Tris pH 6.8, 10 % v/v glycerol, 2 % w/v SDS, bromophenol blue
- SDS wash buffer: 100 mM Tris-HCl pH 7.5, 1 % w/v SDS, 10 mM DTT
- Solution I: 25 mM Tris-HCl pH 8, 10 mM EDTA, 50 mM Glucose +6 μ l/ml RNase A
- Solution II: 0.2 M NaOH, 1 % SDS
- Solution III: 3 M KAc pH4.8
- 5X SSC pH 7.4: 0.75 M NaCl, 0.075 M Na_3 .Citrate.2 H_2O
- T7 cold transcription buffer: 40 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, 10 mM NaCl, 2 mM spermidine-HCl
- T7 hot transcription buffer: 40 mM Tris pH 7.5, 6 mM $MgCl_2$, 10 mM NaCl, 2 mM spermidine-HCl
- 1X TBE: 89 mM Tris base, 89 mM boric acid, 2 mM EDTA
- 1X TBS, pH 7.4: 25 mM Tris base, 150 mM NaCl, 2 mM KCl

- TG loading buffer: 50 mM Tris base, 50 mM glycine, 40 % v/v glycerol, bromophenol blue, xylene cyanol
- TG running buffer: 50mM Tris base, 50mM glycine

**Chapter 3. Investigations into aspects of splicing: splicing optimization,
complex analysis, studies of U1 snRNP and DNase**

3.1 Optimization of splicing

3.2 Characterization of spliceosomal complexes by native agarose gel electrophoresis

3.3 A study of the complexes formed by fluorescent fusion proteins in NE

3.4 An unusual effect of U1 snRNA knockdown by DNA-mediated RNase H cleavage

3.5 The effect of DNase on splicing

Introduction

In order to ask detailed questions about splicing using the TIRF microscopy techniques recently developed in the laboratory it was necessary to identify and validate a model system. Splicing reactions would be used to analyse the factors bound to RNA during splicing on the single molecule level. In a splicing reaction, an in-vitro synthesised pre-mRNA is incubated in NE in conditions resembling those found in the nucleus. Over time the pre-mRNA is converted into the products as described in Figure 1.1B. Nuclear extract contains all the necessary factors for splicing but splicing efficiency will vary greatly between different RNAs. When working at the single molecule level only a very small proportion of the total RNAs within the reaction will be analysed. In order to detect high levels of co-localisation between a factor and an RNA it is necessary to use an RNA which splices with high efficiency so that the efficiency of spliceosome formation will be maximised.

The majority of experiments were carried out using transcripts derived from the rabbit β -globin pre-mRNA. The rabbit β -globin pre-mRNA transcript consists of the second exon, a shortened second intron and part of the third exon of the rabbit β -globin pre-mRNA (Skordis et al., 2003). This pre-mRNA is spliced with high efficiency in HeLa nuclear extract, for example see Figure 3.1A. Radioactive nucleotides are incorporated into the pre-mRNA to allow detection of the products of a splicing reaction after electrophoresis, using phosphor screens. The pre-mRNA transcript was transcribed from PCR products which had been generated using either the oligonucleotides P16 and P17, or P16 and β G-E3no5'ss (Section 2.1.1). The transcript made from the PCR fragment generated using P16 and β G-E3no5'ss was 388 nt in length. Splicing of this

pre-mRNA yields an mRNA of 282 nt and a lariat of 106 nt; the intermediate 5' exon and 3'exon lariats are 226 nt and 162 nt in length respectively. Those RNAs transcribed from PCR fragments generated using P16 and P17 were longer by 9 nt due to the presence of a consensus 5' SS at the 3' end of the transcript. This was designed to enhance stability of the RNA. The products, were assigned based upon previous observations in the laboratory together with their migration relative to radioactive markers and their dependence on the time of incubation.

Experiments were carried out to optimize the splicing reaction, to define the spliceosomal complexes formed on the β -globin RNA, and to stall the spliceosome at specific points during formation.

3.1 Optimization of splicing

A series of different components of the splicing reaction were altered to assay their effect on the efficiency of the splicing reaction.

Potassium Source

Reichert and Moore, (2000) reported that replacing the KCl in a splicing reaction with potassium acetate or potassium glutamate increased the efficiency of splicing because it counteracted the inhibitory effect of the chloride ions. To investigate the effect of KCl and potassium glutamate on β -globin splicing a standard splicing reaction was set up (Section 2.1.3) in which the 50 mM potassium glutamate was replaced with 50 mM KCl. There was very little difference in splicing efficiency under these conditions (Figure 3.1A). It was concluded that the source of potassium was not of especial importance for the efficiency of β -globin splicing.

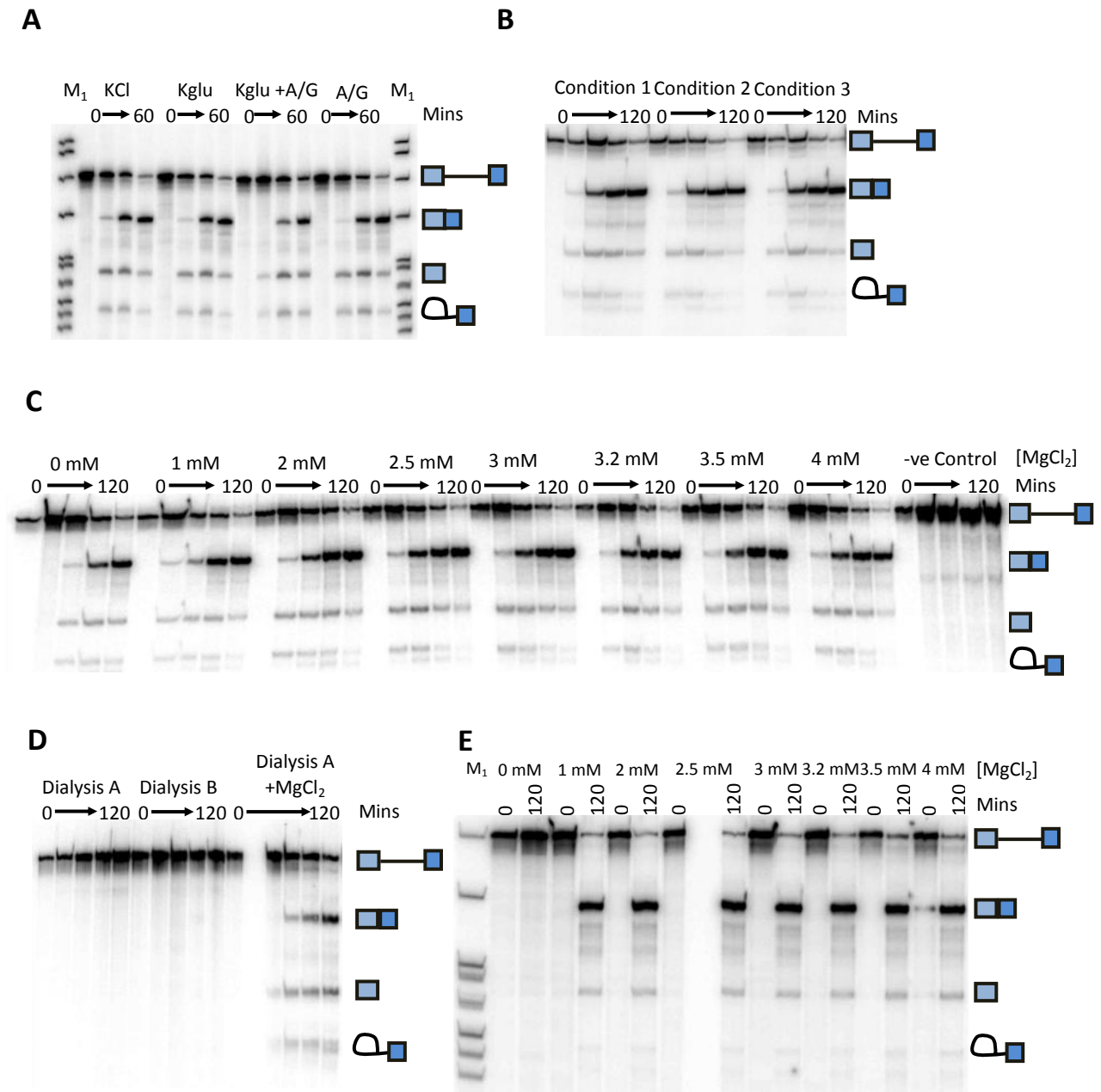


Figure 3.1: Optimization of splicing reactions. **(A)** Testing of different sources for additional potassium and the effect of arginine and glutamate on splicing. A/G represents the addition of arginine and glutamate. The box and line illustrations by the right hand side of the figure represent the precursor, intermediates and products of splicing as shown in Figure 1.1B. Time points of 0, 15, 30 and 60 minutes were taken. M₁ refers to the marker (Table 2.2). **(B)** Testing the requirements of a splicing reaction: Condition 1 is a standard splicing reaction with the addition of 20 mM Hepes, 0.13 % NP40 and 6 % RNasin; Condition 2 is a standard splicing reaction and Condition 3 is a standard splicing reaction with 3.2 mM Kglutamate instead of 50 mM. Time points of 0, 15, 30, 60 and 120 minutes were taken. **(C)** Titration of MgCl₂ into a splicing reaction made with commercial NE. Time points of 0, 15, 30, 60 and 120 minutes were taken. **(D)** Dialysed 'home-made' NE unable to splice in 0mM MgCl₂. Time points of 0, 15, 30, 60 and 120 minutes were taken. Dialysis A refers to the technique described in Section 2.2.3, Dialysis B refers to dialysis using fine tubing. **(E)** Titration of MgCl₂ into a splicing reaction made with commercial NE re-dialysed using Dialysis A.

Addition of Arginine and Glutamic acid

Some splicing factors have been shown to have higher solubility in 50 mM L-arginine and 50 mM L-glutamic acid (Golovanov et al., 2004). The effect of adding both to a splicing reaction was tested because this enhanced solubility may have reduced the amount of precipitated material in the NE and perhaps thereby enhanced splicing. The L-arginine and L-glutamic acid were added to the standard splicing reaction (Section 2.1.3) either along with the potassium glutamate solution or instead of it (both solutions were pH~7.5). Surprisingly splicing under both conditions was of very similar efficiency and was also of similar efficiency when compared to potassium glutamate addition (Figure 3.1A). This was surprising, as it was expected that the sample lacking additional potassium glutamate would splice with a lower efficiency; the observation that it has not, suggests that the 50 mM KCl, which is added to the splicing reaction as part of the NE, is sufficient to support efficient splicing of the β -globin pre-mRNA. This is supported by the finding that splicing efficiency is not altered when the concentration of potassium glutamate added to a splicing reaction is reduced from 50 mM to 3.2 mM (Figure 3.1B). However since these results were obtained using the β -globin RNA and other RNAs might differ in their Potassium glutamate requirements the addition of 50 mM potassium glutamate was kept as part of the standard splicing reaction.

Magnesium chloride concentration

The concentration of $MgCl_2$ in a splicing reaction was also investigated. Titrations of $MgCl_2$ were carried out in reactions using commercial HeLa NE and the β -globin RNA (Figure 3.1C). The concentration of $MgCl_2$ in the splicing reaction appeared to have

very little effect on the efficiency of splicing with only the lower concentrations showing any loss of efficiency. Surprisingly, in these experiments the sample without additional MgCl_2 was able to splice.

To test whether this was due to the source of the NE, a 'home-made' HEK-293T cell NE was also tested for splicing in a reaction lacking MgCl_2 . The last stage in the preparation of NE is to dialyse against Buffer D. Two separate NE dialysis techniques were tested, the technique described in section 2.2.3 and dialysis using fine tubing. These were tested in case one of the techniques led to MgCl_2 coming through the extraction process from the cells. Splicing in the absence of MgCl_2 was not observed using either dialysis technique but was observed in the presence of 3.2 mM MgCl_2 (Figure 3.1D).

In light of this result, to investigate whether commercial NE contained MgCl_2 , commercial NE was re-dialysed against standard Buffer D using the method described in section 2.2.3. Splicing with the re-dialysed commercial NE showed that it did indeed contain MgCl_2 , as splicing in the absence of added MgCl_2 was not observed (Figure 3.1E). There was still very little difference in splicing efficiency when the MgCl_2 concentration was between 1 mM and 4 mM (Figure 3.1E). In conclusion, splicing of the β -globin pre-mRNA required MgCl_2 but is not increased by increasing the concentration above 1 mM. The lab standard of 3.2 mM MgCl_2 was kept for further experiments.

Other experiments looked at the effect of varying the ATP and Phosphocreatine concentrations on splicing efficiency, but these showed no effect on splicing efficiency over the concentrations tested (data not shown). Addition of 20 mM Hepes pH7.5,

0.13 % NP-40 and 6 % RNasin were also tested and showed no enhancement or inhibition of splicing (Figure 3.1B).

3.2 Characterization of spliceosomal complexes by native agarose gel electrophoresis

As described in section 1.1 the spliceosome forms on a pre-mRNA through a sequential series of complexes; these complexes can be resolved by native LMP agarose gel electrophoresis (Das and Reed, 1999). In order to study the spliceosome at the single molecule level, we planned to use inhibitors to stop spliceosome formation at certain points. To be able to do this the complexes formed on β -globin RNA needed to be defined. The different spliceosomal complexes that can be detected by native agarose gel electrophoresis are described in Table 3.1; the different spliceosomal complexes have different requirements and can be resolved on different percentage LMP agarose gels.

Complex	H	E	A	B	C
ATP requirement	-	-	+	+	+
snRNPs involved	-	U1/U2	U1/U2	U1/U2/U4/ U5/U6	U2/U5/U6
Heparin sensitivity	insensitive	sensitive	insensitive	insensitive	insensitive
% LMP agarose required for resolution	1.5/2	1.5	2	2	2

Table 3.1: Spliceosomal complexes and their requirements.

Assignment of E complex

The first detectable and specific complex formed on pre-mRNA is E complex; E complex is defined as being ATP-independent and heparin-sensitive; it can be resolved on a native 1.5 % LMP agarose gel (Das and Reed, 1999). By depleting the NE of ATP before addition to a standard splicing reaction, which lacks ATP and Phosphocreatine,

spliceosome formation is stalled at E complex. In previous work (Das and Reed, 1999) the spliceosome was stalled at E complex by pre-incubation of the NE at room temperature for 30 minutes. When attempts were made to repeat this procedure and the complexes formed were resolved on a native 2 % LMP agarose gel in the presence of heparin, a small heparin-resistant complex was detected (Figure 3.2A).

It was hypothesised that this small complex may represent low levels of A complex and that it may be due to incomplete depletion of ATP. To deplete ATP fully, 10 U/ml of hexokinase (Sigma) and 0.4 mg/ml glucose were added to the NE prior to incubation at 30 °C for 30 minutes (Tatei et al., 1989). Depletion of ATP was monitored by thin layer chromatography analysis of exogenously added [γ -³²P] ATP (Figure 3.2B). This showed that the addition of hexokinase and glucose led to the complete removal of ATP (Figure 3.2B). Electrohoresis of samples on a native 2 % LMP agarose gel showed that no heparin-resistant complexes had formed (Figure 3.2C). Analysis on a native 1.5 % LMP agarose gel in the absence of heparin showed that all the RNA remained in the wells on electrophoresis and no discrete complexes were resolved. (Figure 3.3A). This showed that ATP-depletion using hexokinase and glucose was an unsuitable method for stalling assembly at E complex.

To characterise E complex using a different method snRNAs in the NE were digested using DNA oligonucleotide-mediated RNase H cleavage. Digestion of the 5' end of U1 snRNA using the Anti-U1 oligonucleotide (Table 3.2) at 8 pmol/ μ l led to an almost complete loss of complex formation (Figure 3.3A); this would be expected as E complex is dependent on U1 snRNP (Reed, 1990). The low levels of complex formation could be explained by the low levels of full length U1 snRNA left in the extract after

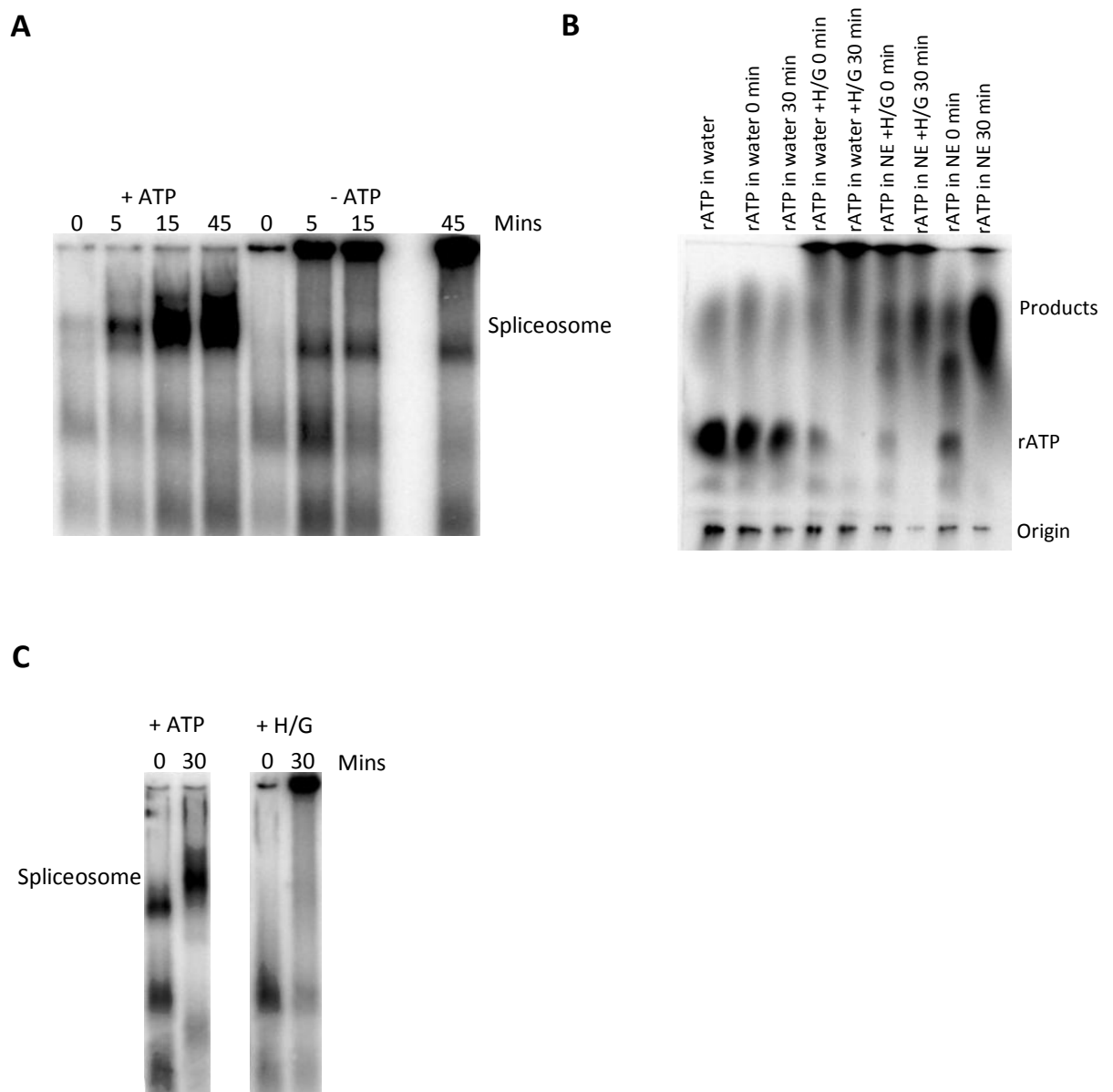


Figure 3.2: ATP depletion mediated by Hexokinase and glucose. (A) Formation of a small heparin-resistant complex of similar size to A complex under standard E complex conditions on a 2 % LMP agarose gel. **(B)** Image of a thin layer chromatography plate showing optimization of ATP depletion using Hexokinase and glucose where H/G stands for the addition of hexokinase and glucose. **(C)** A 2 % LMP agarose gel showing no formation of heparin resistant spliceosomal complexes upon depletion of ATP using Hexokinase/glucose.

Name	Sequence
Anti-U1	TGCCAGGTAAGTAT
E15	AGGCCGAGAAGCGAT
L15	CAGATACTACAGTTG
U4B	GGAAAAGTTTTCAATTAG
U6A	GCCATGCTAATCTTCTCT
U6B	CGCTTCACGAATTTGCGT
U1 19-33	TGATCATGGTATCTC
U1 50-65	GGATAAGCCTCGCCC
U1 90-105	TTTGGGGAAATCGCA
U1 120-135	ACCACAAATTATGCA
Random	NNNNNNNNNNNNNNNN
Pos 1	GCCTTCAGGTTAGG
Pos 2	GAACAATCAAGGGT

Table 3.2: Sequences of oligonucleotides used for RNase H mediated degradation.

digestion (Figure 3.3B). Nucleotides modified at the 2' position of the ribose ring with a methyl group prevent RNase H activity, so oligonucleotides containing this modification and complementarity to the 5' terminus of U1 snRNA (α -U1; Table 2.1) form stable hybrids. Pre-incubation of the NE with 8 pmol/ μ l of α -U1 oligonucleotide for 15 minutes, led to loss of complex formation and an accumulation of H complex (Figure 3.3A). This showed that the low level of complex formation observed in samples where the U1 snRNA was digested using RNase H are probably due to residual full-length U1 snRNA.

Although U2 snRNP is required for E complex formation (Donmez et al., 2004) it does not associate with the branch-point until A complex. Therefore digestion of U2 snRNA may allow us to stall the spliceosome at E complex. Digestion of the 5' end of U2 snRNA using the E15 oligonucleotide (Table 3.2; Black et al., 1985) at a concentration of 2 pmol/ μ l led to slower complex formation (Figure 3.3A). However this treatment did not block splicing (Figure 3.3C); this suggests either that the digestion was not complete or that U2 snRNA lacking its 5' terminus is able to progress through the

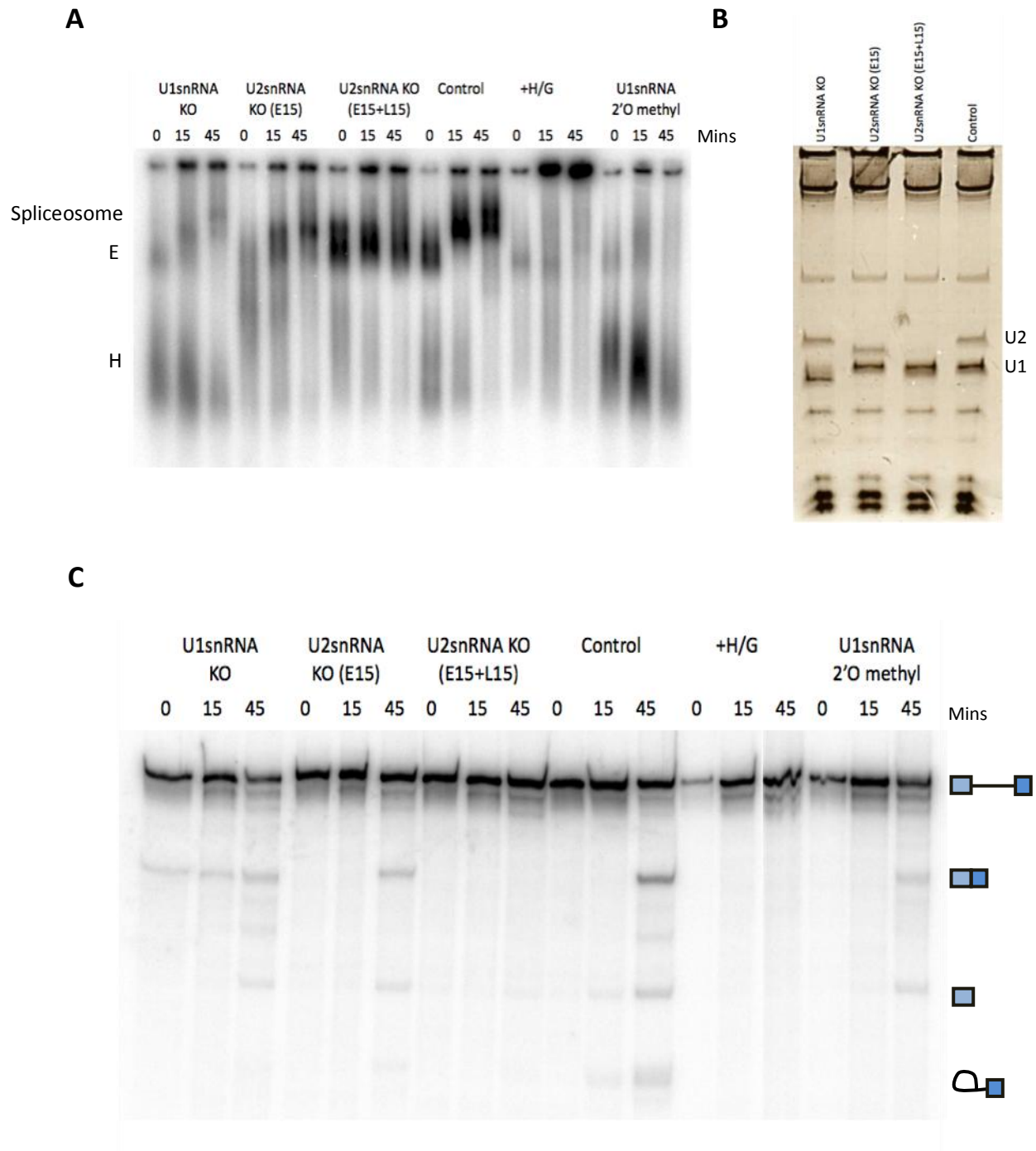


Figure 3.3: Assignment of possible E complex using snRNA degradation. (A) Formation of spliceosomal complexes in the absence of heparin on a 1.5 % LMP agarose gel under different conditions. Complexes were inhibited either by knockout (KO) of specific snRNAs using RNase H digestion (U1 snRNA KO, U2 snRNA KO (E15) and U2 snRNA KO (E15+L15), or with a 2'O methyl oligonucleotide complementary to the 5' end of U1 snRNA (U1snRNA 2'O methyl) or through treatment with Hexokinase/glucose to deplete ATP (+H/G). Band E represents the assignment of a possible E complex. **(B)** Ethidium-stained 6 % denaturing urea polyacrylamide gel showing the snRNA present in the samples used for studies of complex formation in Figure 3.6A. **(C)** RNA products formed in the splicing reactions used in Figure 3.6A run on a denaturing urea 6 % polyacrylamide gel.

spliceosome, albeit at a slower rate. Digestion of the U2 snRNA using the E15 oligonucleotide in conjunction with the L15 oligonucleotide (Table 3.2; Black et al., 1985) at a concentration of 2 pmol/μl led to loss of splicing (Figure 3.3C) and efficient digestion of U2 snRNA (Figure 3.3B). The spliceosome formed in this extract was stalled at a complex smaller than those formed in control experiments (Figure 3.3A). Based on this complex's requirement for intact U1 snRNA but not U2 snRNA together with the fact that it was not present in heparin-treated samples run on a native 2 % LMP agarose gel, (Figure 3.6A) it was labelled E complex. This suggested that the complex observed in samples pre-incubated at room temperature for 30 minutes was a low level of A complex being formed due to incomplete ATP depletion.

Many of the experiments described in the later chapters of this thesis were carried out on a variant of the β-globin RNA, which contained a consensus 5' SS i.e. globin C. Globin C was constructed by replacing the β-globin 5' SS sequence with the consensus 5' SS sequence (CAGGUAAGU) by site-directed mutagenesis using the oligonucleotides Fpri-Cons-EP and Rpri-Cons-EP (Table 3.3). The formation of E complex on this pre-mRNA was assessed in the following experiments. Firstly, E complex could be detected in NE that had been pre-incubation at 30 °C for 30 minutes (Figure 3.4A) in the absence of hexokinase and glucose.

Name	Sequence
Fpri-Cons-EP	CGTGGATCCTGAGAACTTCAGGTAAGTTTGGGGACCCTTG
Rpri-Cons-EP	CAAGGGTCCCCAACTTACCTGAAGTTCTCAGGATCCACG
GGmutF	TGCCTTCTTCTTTTTCCTACGGCTCCTGGGCAACGTGCTGG
GGmutR	CCAGCACGTTGCCAGGAGCCGTAGGAAAAAGAAGAAGGCA
MutF	GATCCTGAGAACTTCAGGATTGGGGACCCTTGATTGT
MutR	ACATCAAGGGTCCCCAATCCTGAAGTTCTCAGGATC

Table 3.3: Sequences of oligonucleotides used for site directed mutagenesis.

The complex formed on this pre-mRNA when ATP was depleted was dependent on the presence of U1 snRNA. This was shown by RNase H digestion of the 5' end of U1 snRNA using the anti U1 oligonucleotide at 8 pmol/ μ l which led to loss of the complex (Figure 3.4A). This complex also did not form on a second variant of the β -globin RNA, globin M (Figure 3.4A), in which the 5' SS had been deleted by site-directed mutagenesis with the MutF and MutR primers (Table 3.3) to leave the sequence CGGAU.

The complex assembled in NE depleted of ATP by incubation at 30 $^{\circ}$ C for 30 minutes was also sensitive to heparin; it was not detected when heparin was added and the sample was run on a native 2 % LMP agarose gel (Figure 3.4B). The lack of a requirement for the addition of hexokinase and glucose was supported by thin layer chromatography experiments in which incubation of NE at 30 $^{\circ}$ C for 30 minutes led to conversion of exogenous [γ - 32 P] ATP to unidentified products (Figure 3.2B). On the basis of U1-dependence and heparin sensitivity, this complex was assigned as E complex.

When ATP is depleted from an NE to stall spliceosome formation at E complex a large proportion of the RNA in a sample will not move into the gel and instead will remain in the wells (for example, Figures 3.4A and 3.4B). To test whether this was due to dephosphorylation of spliceosomal factors, experiments with the Phos-stop phosphatase inhibitor cocktail (Roche) were carried out. Addition of Phos-stop, at 2x the manufacturer's recommended concentration prior to incubation at 30 $^{\circ}$ C for 30 minutes, led to far less RNA remaining in the wells leading to an apparent increase in the efficiency of E complex formation (Figure 3.4C). These results were consistent

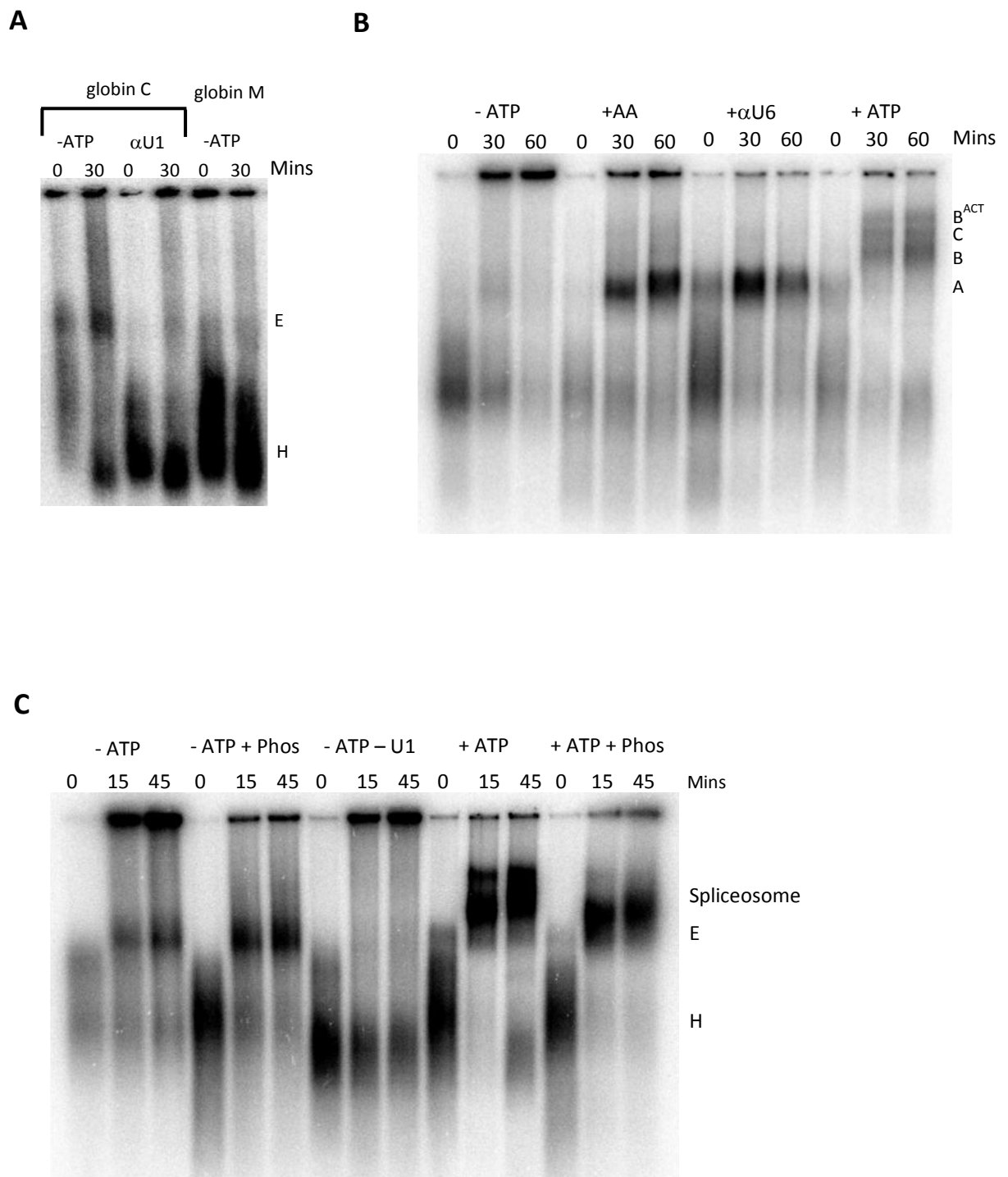


Figure 3.4: E complex assembly and the effect of phosphatase inhibition. **(A)** Depletion of ATP by incubation at 30 °C for 30 minutes showing E complex formation in the absence of Hexokinase and glucose, α-U1 indicates U1 depletion by RNase H mediated snRNA degradation. **(B)** Depletion of ATP by incubation at 30 °C for 30 minutes showing a loss of heparin resistant complexes. As markers for A complex samples +AA and +U6 were treated with either 0.3 mM anacardic acid (AA) or 1 μM α-U6 oligonucleotide (+αU6; Table 2.1). **(C)** Addition of Phos-stop phosphatase inhibitor cocktail (Roche) leads to less RNA staying in the wells. Samples labelled +Phos have had phosphatase inhibitor added, samples labelled -U1 have been depleted of U1 snRNA by incubation with the α-U1 2'O methyl oligonucleotide.

with the finding that addition of PP1 phosphatase to NE inhibited E complex formation (Mermoud et al., 1994) and suggest that protein dephosphorylation may account for some of the difficulties in detecting E complex. To allow addition of Phos-stop to the splicing reaction this experiment was carried out using 45 % NE instead of 50 %.

Assignment of the A, B and C complexes

The other complexes that can be resolved by native LMP agarose gel electrophoresis are the A,B and C complexes; these are ATP-dependent, heparin-insensitive and can be resolved on native 2 % LMP agarose gels. A complex was assigned using histone acetyl transferase inhibitors, anacardic acid and garcinol that have been shown to stall splicing at A complex (Kuhn et al., 2009). Nuclear extract was incubated at 30 °C for 15 minutes with anacardic acid, garcinol or DMSO prior to assembly of the splicing reaction. The control was incubated with DMSO as both Anacardic acid and Garcinol were dissolved in it. Resolution of heparin-resistant complexes on a native 2 % LMP agarose gel showed that both anacardic acid and garcinol stalled the spliceosome at a complex that was smaller than the major complex formed on the β -globin RNA (Figures 3.5A and B). In the absence of anacardic acid or garcinol this complex was only observed in the earliest time points and was assigned as A complex. There were only two other bands observed on native agarose gels with the β -globin RNA and these were assigned as B and C complexes based on their relative size and appearance (Figure 3.5A, DMSO sample). The assignment of A complex was confirmed when assembly was stalled, on the globin C pre-mRNA, after incubation of the NE with a 2'O methyl oligonucleotide complementary to the U6 snRNA (Table 2.1; Donmez et al., 2007; Figure 3.5C). Pre-incubation was for 15 minutes at 30°C.

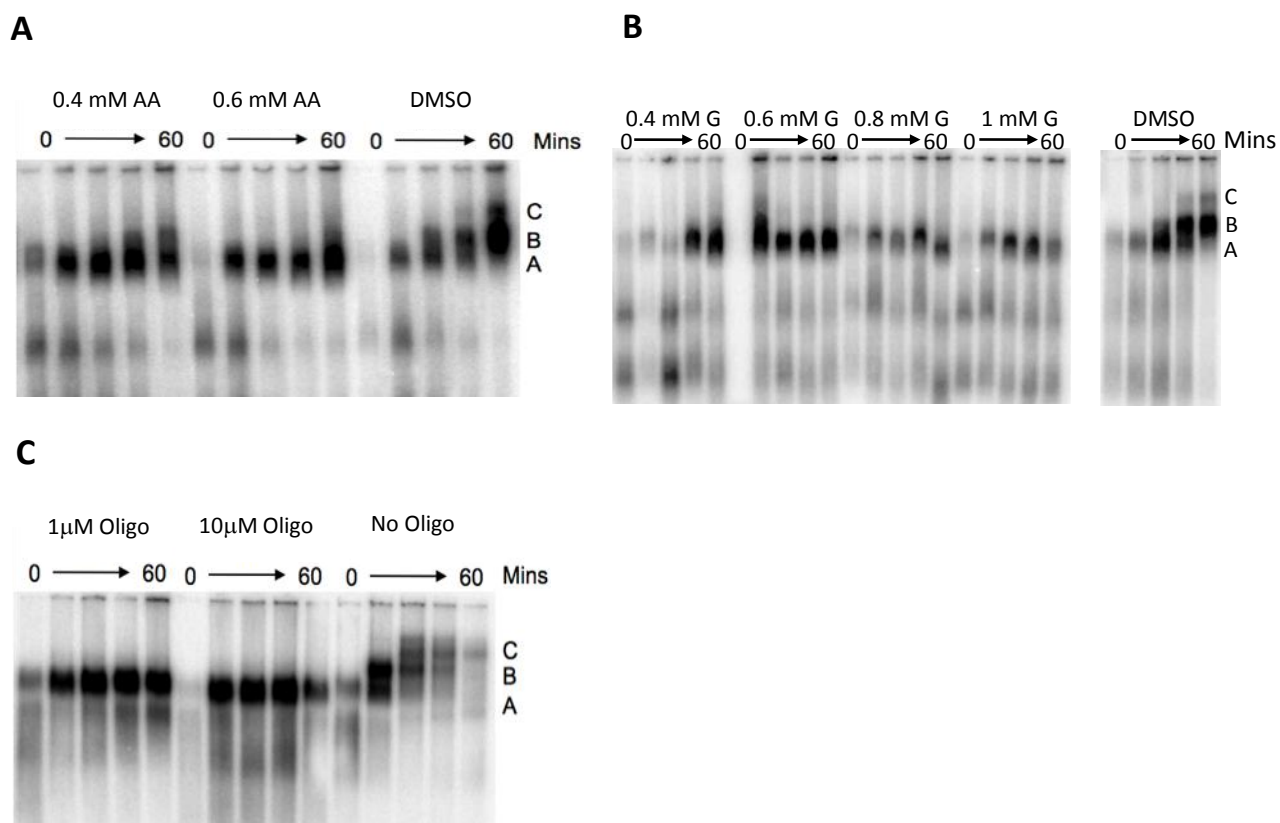


Figure 3.5: Characterization of A complex. **(A)** Stalling spliceosome formation at complex A using anacardic acid (AA) on the β -globin RNA where time points were 0, 1, 5, 15 and 60 minutes; DMSO was used as a control as AA is dissolved in DMSO. The letters on the right hand side of the panel are the assignments of the spliceosomal complexes. **(B)** Stalling spliceosome formation at Complex A using garcinol (G) on the β -globin RNA where time points were 0, 1, 5, 15 and 60 minutes. **(C)** Stalling spliceosome formation at A complex on the globin C RNA using the α -U6 oligonucleotide (Table 2.1) where time points were 0, 5, 15, 30 and 60 minutes.

To investigate the assignments of A, B and C complexes further snRNA digestions were carried out. A complex and the larger B and C complexes should be dependent on the U1 and U2 snRNAs, whilst B and C complexes should also be dependent on the U4, U5 and U6 snRNAs. All of the larger heparin-resistant complexes were lost following digestion or sequestration of U1 snRNA or digestion of U2 snRNA, consistent with expectations (Figure 3.6A). Digestion of U4 snRNA and U6 snRNA should lead to an accumulation of A complex and a loss of B and C complexes. However, digestion of U4 snRNA using the U4B oligonucleotide (Table 3.2; Black and Steitz, 1986) at 2 pmol/ μ l led to an almost complete loss of complex formation with a very small amount of complex having accumulated by 45 minutes (Figure 3.6A). Conversely digestion of the U6 snRNA using the U6A and U6B oligonucleotides (Table 3.2; Black and Steitz, 1986) together had little effect on spliceosomal complex assembly (Figure 3.6A). This was despite the treatments to digest U4 snRNA and U6 snRNA having similar effects on splicing efficiency and the observed levels of the snRNA (Figure 3.6B and C). Since the oligonucleotides would cause digestion of the complementary sequences that mediate base-pairing between U4 and U6 snRNA, their effects should be similar. The difference is likely to be due to incomplete digestion of the U6 snRNA.

Analysis of the complexes formed on globin C pre-mRNA appeared to show the formation of four heparin-resistant complexes when run on a native 2 % LMP agarose gel (for example see Figure 3.5C). These were originally assigned as the A, B, C spliceosomal complexes and a post spliceosomal complex (PSC), with the PSC migrating between the putative B and C complexes. However, analysis of the splicing

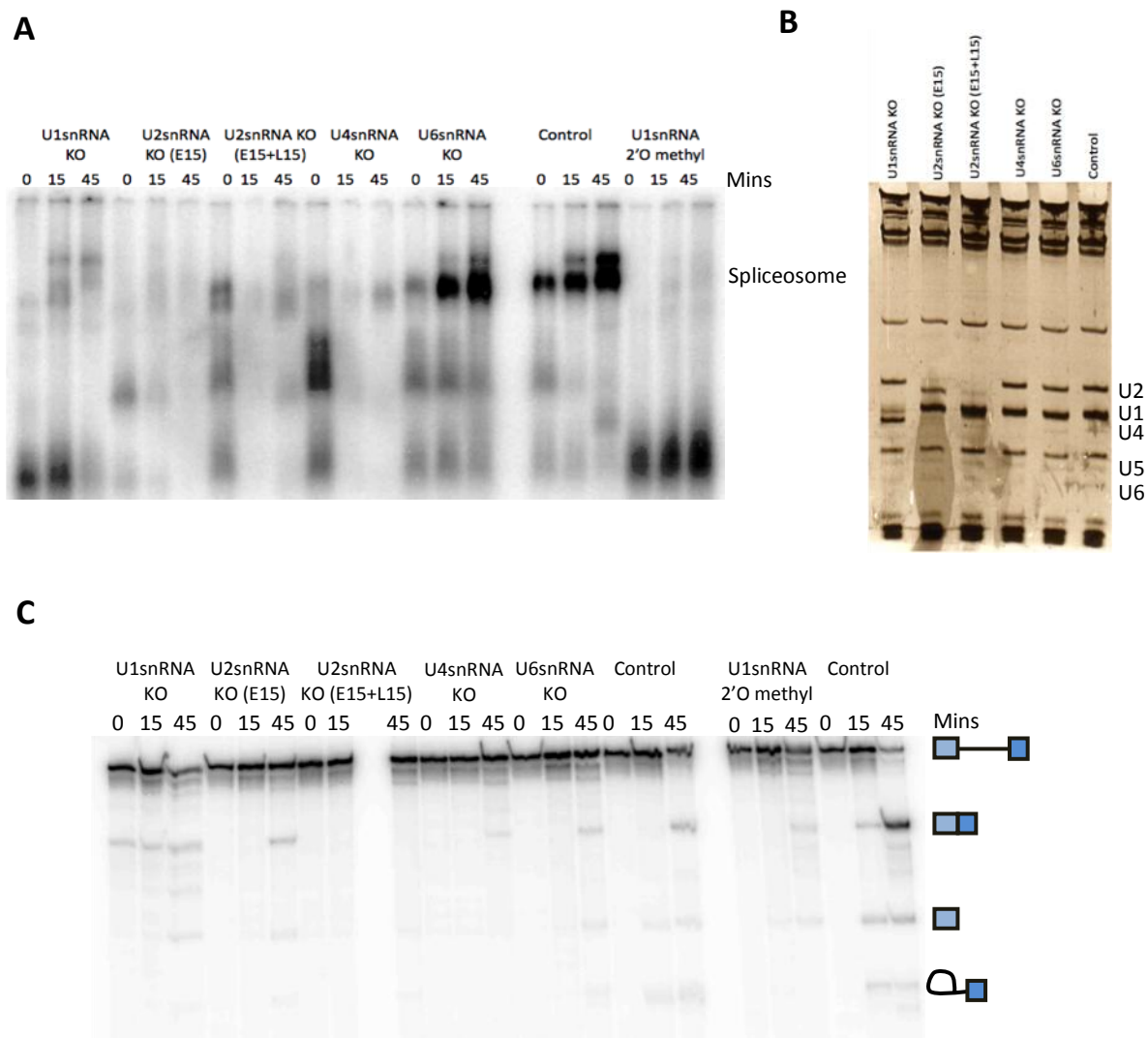


Figure 3.6: Characterization of the spliceosome using RNase H-mediated snRNA degradation.
(A) LMP agarose gel (2 %) showing the heparin-resistant complexes formed when the snRNAs, U1, U2, U4 or U6 were knocked out by RNase H-mediated degradation. **(B)** Ethidium-stained denaturing urea 6 % polyacrylamide gel showing degradation of snRNA in the NE used for splicing shown in Figure 3.6A. **(C)** Splicing products formed in reactions used in Figure 3.6A run on a denaturing urea 6 % polyacrylamide gel.

products and intermediates being formed during a time course in relation to the complex formation, showed that this could not be the case as splicing intermediates only appeared once the putative PSC had assembled (Figure 3.7A, compare panels one and two); this was more consistent with the putative PSC being C complex. This observation raised the possibility that what was observed was actually the formation of a new splicing complex.

To test this, a derivative of globin C, globin C-GG was made, in which the AG dinucleotide at the 3' SS was replaced with GG. This mutation leads to the second step of splicing being blocked but allows the first step to occur on AG-independent introns (Figure 3.7C; Gozani et al., 1994, Jurica et al., 2002). globin C-GG was made by site-directed mutagenesis of globin C using the GGmutF and GGmutR oligonucleotides (Table 3.3). The shortened 3' exon-lariat in Figure 3.7C was hypothesised to be due to degradation of the transcript after release from the spliceosome. As the first step of splicing occurs between the B* and C complexes and the second step occurs after C complex formation, the globin C-GG RNA should form C complex but not any PSC. Therefore, if the fourth complex to form is a PSC it should not form on the Globin C-GG RNA. Analysis of the complexes showed that they all formed on globin C-GG, so it was inferred that there are four heparin-resistant complexes. Whilst this was in progress, a new complex, B^{ACT}, was identified that migrated more slowly than C complex (Bessanov et al., 2010). I infer that the complexes shown in Figure 3.7B are complexes A, B, B^{ACT} and C.

Interestingly, when investigating the effect of phosphatase inhibition by Phos-stop on overall splicing efficiency, the inhibitor was also observed to influence the migration

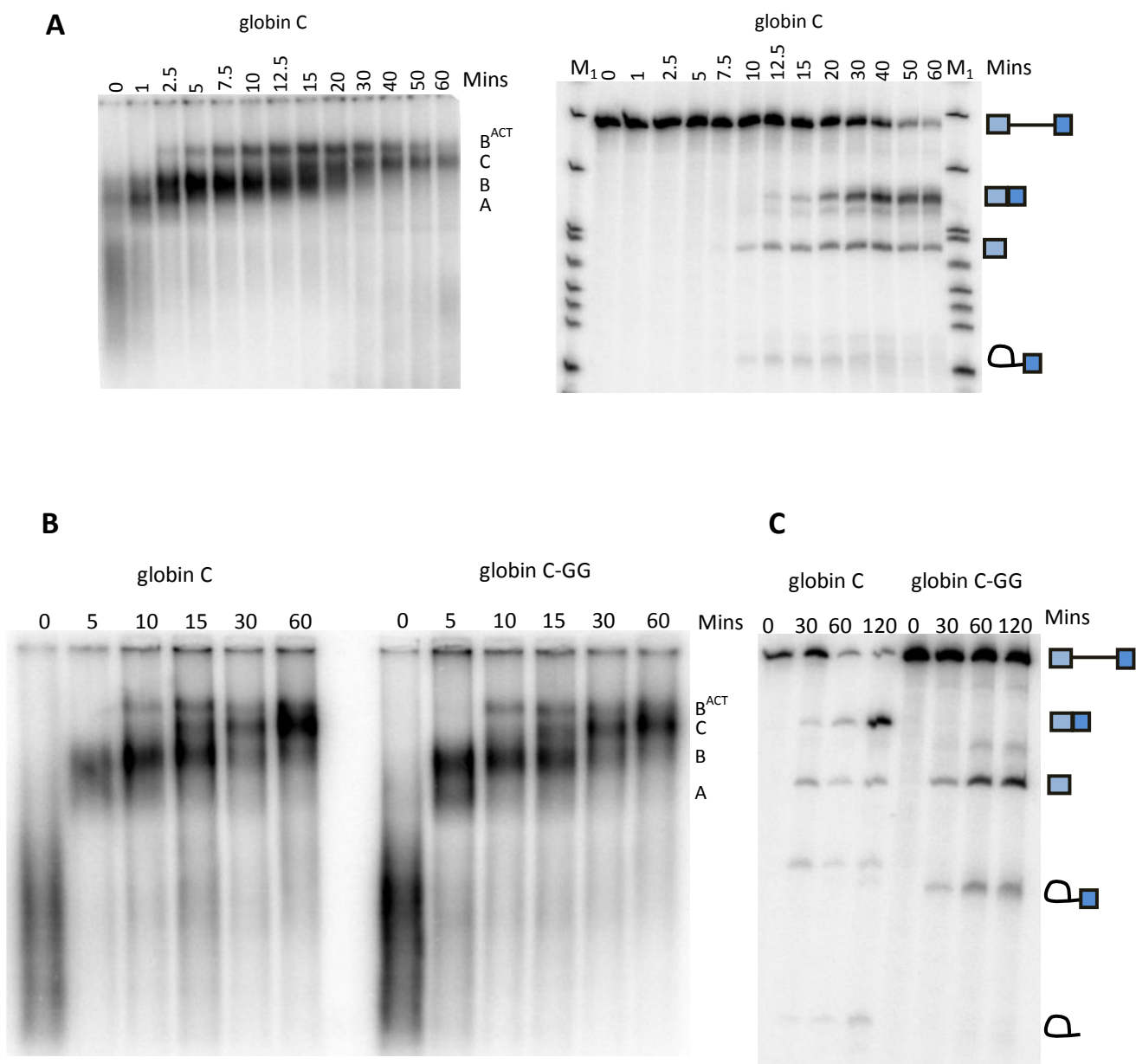


Figure 3.7: Confirmation of the presence of B^{ACT} and C complexes. **(A)** Time course of complex formation shown on a 2 % LMP agarose gel with the concomitant splicing, run on a denaturing urea 6 % polyacrylamide gel. **(B)** A 2 % LMP agarose gel of the heparin-resistant complexes formed on the globin C and globin C-GG RNAs. **(C)** A denaturing urea 6 % polyacrylamide gel of the splicing products formed in the reactions used in Figure 3.7B.

patterns of the A, B, B^{ACT} and C complexes. Nuclear extract that had been incubated at 30 °C for 30 minutes, in the presence of Phos-stop, was added to a standard splicing reaction and analysed by native agarose gel electrophoresis in the absence of heparin. The results shown in Figure 3.8A show that, although splicing had not been completely inhibited, complex formation appeared to be inhibited and the complexes formed upon incubation with different concentrations of Phos-stop were different sizes. It appeared that phosphatase inhibition caused the complexes to migrate through the gel further. This may be because the inhibition of phosphatases led to an overall increase in the negative charge of the complex, or because the inhibition led to the removal of proteins that are associated only when dephosphorylated. This effect was lost when the samples were treated with heparin before being analysed by native agarose gel electrophoresis, although B^{ACT} complex appeared to be destabilised as it was not detected when heparin was added to phosphatase-treated samples (Figure 3.8B). The observation that this effect is lost in the presence of heparin explains why it has not been observed before (Mermoud et al., 1992; Shi et al., 2006) and may partially explain why splicing is inhibited by inhibition of, or depletion of phosphatases.

3.3 A study of the complexes formed by fluorescent fusion proteins in NE

As we planned to use fluorescent fusion proteins to study the spliceosome at the single molecule level, it was important to determine whether or not the fluorescent fusion proteins to be used were functional. One method that was used to assess this was native agarose gel electrophoresis. The hypothesis that led to this avenue of research was that over-expressed fluorescent fusion proteins in NE should be part of the spliceosomal complexes formed on exogenously added RNA.

Native gel analysis of NE expressing a fluorescently-labelled version of the U1 snRNP protein U1-A (mEGFP-U1-A), in the absence of pre-mRNA and heparin showed U1-A to migrate in three distinct bands (Figure 3.9A; No knockdown sample). To determine whether the largest of the three bands was the spliceosome formed on endogenously produced pre-mRNAs, a series of experiments were performed. Firstly the effect of U2snRNA knockdown and exogenously added β -globin pre-mRNA were assayed. Nuclear extract containing mEGFP-U1-A was treated with the E15 and L15 oligonucleotides (Section 3.2), to knockdown the U2 snRNA and stall complex formation at E complex. To this NE, an increasing amount of β -globin pre-mRNA was then added. The U2 snRNA knockdown had no effect on the migration of mEGFP-U1-A, but addition of increasing amount of β -globin pre-mRNA led to a loss of the largest complex formed by mEGFP-U1-A and the production of a smear further down the gel (Figure 3.9A).

The complexes observed by native agarose gel electrophoresis of mEGFP-U1-A containing 293T NE were then further characterised. Samples were prepared in which (a) ATP was depleted, (b) U1 snRNA was degraded by RNase H and the Anti-U1 oligonucleotide, (c) U1 snRNA was sequestered using the 2'-O-methyl α -U1 oligonucleotide or (d) treated with heparin prior to loading into the gel. Treatments a,b and c did not affect the migration of the largest complex (Figure 3.9B). But incubation with the 2'-O methyl-modified oligonucleotide led to the appearance of a new band which was slightly smaller than the largest complex; it is possible that this is the U1 snRNP bound to the α -U1 oligonucleotide (Figure 3.9B). The only treatment

that did have an effect on the migration of the largest band was treatment with heparin before native agarose gel electrophoresis (Figure 3.9B). Similar-sized complexes were observed to form when other spliceosomal fluorescent fusion proteins such as mEGFP-SF2 (SFRS-1) were analysed by native agarose gel electrophoresis (Figure 3.9C). The largest complex formed in mEGFP-U1-A containing NE also migrated at a similar position to spliceosomal complexes formed on exogenously-added β -globin RNA (Figure 3.9D).

To determine, conclusively, whether the large complexes observed in mEGFP-U1-A containing NE was the spliceosome or possibly even the supraspliceosome, an amalgamation of four spliceosomes in one huge complex which migrates at 200S on sucrose gradients (Muller et al., 1998, Cohen-Krausz et al., 2007), the position of the individual snRNAs were assessed by Northern blot. This required riboprobes (Blencowe et al., 1989), synthesised using a standard hot transcription from encoding plasmids, which recognised the individual snRNAs specifically (Figure 3.10A).

The results discussed above showing that this complex was dependent on neither U1 snRNA nor U2 snRNA were hypothesised to have been observed for one of two reasons: either the digested snRNAs were being stabilised within the complex via the protein-protein interactions between components of the spliceosome or the snRNAs were not accessible to the RNase H and DNA oligonucleotide and thus were protected.

To determine whether the complexes seen were spliceosomes and/or supraspliceosomes, Northern blot analysis of the position of the snRNAs on a native agarose gel was performed. NE containing mEGFP-U1-A was run into a native 1.5 % LMP agarose gel and imaged using a typhoon scanner and a 488 nm laser (Figure 3.10B

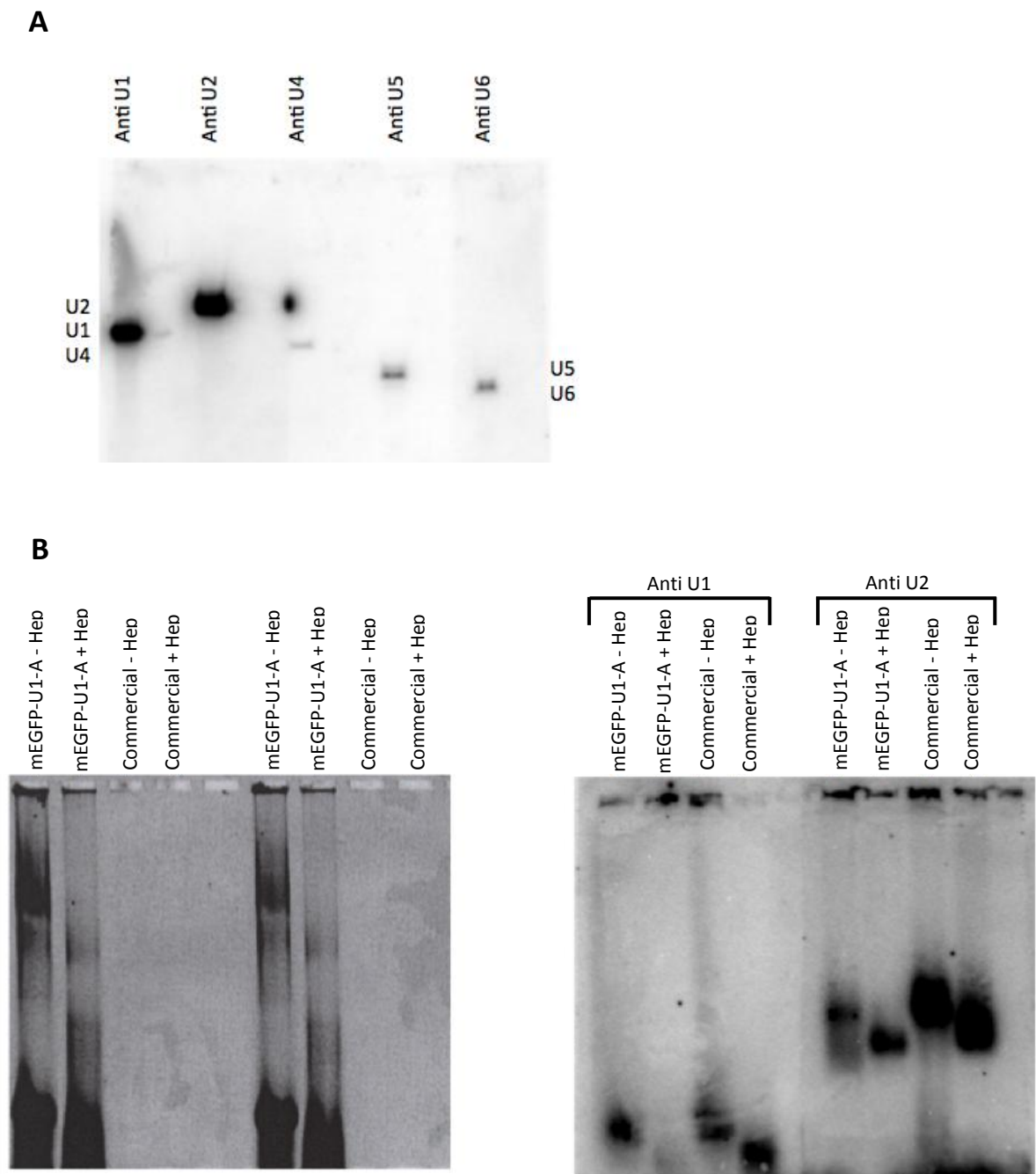


Figure 3.10: Detection of snRNA using Northern blotting. (A) Detection of snRNAs in NE run on a denaturing urea 6 % polyacrylamide gel by Northern blotting using riboprobes against U1, U2, U4, U5 and U6 snRNA (Anti U1 to U6). (B) Determination of the position of the U1 and U2s nRNAs relative to the complexes formed by mEGFP-U1-A. The first panel shows the complexes formed by mEGFP-U1-A, resolved on a 1.5 % LMP agarose gel in the presence (+Hep) or absence (-Hep) of heparin. The second panel shows a Northern blot of the agarose gel shown in the first panel showing the positions of both U1 (Anti U1) and U2 (Anti U2) snRNA.

panel one) before being transferred to NX membrane. Five sets of replicates were run on the same gel to allow probing of the blot with the five riboprobes individually. Once incubated with the riboprobes as described in Section 2.1.7, the membrane was exposed to a phosphor screen. Only the U1 snRNA and U2 snRNA probes produced any signal, probably due to the comparably weak detection of U4, U5 and U6 snRNAs by their respective riboprobes (Figure 3.10A) and therefore only the signal from U1 snRNA and U2 snRNA is displayed in Figure 3.10B. What is clear is that the snRNAs in the NE did not migrate at the same position as the large complex formed by mEGFP-U1-A (Figure 3.10B) and thus the complex is unlikely to be splicing-related and is instead probably an artefact produced by native LMP agarose gel electrophoresis of NE containing overexpressed mEGFP-U1-A.

3.4 An unusual effect of U1 snRNA knockdown by DNA-mediated RNase H cleavage

Whilst investigating the effects of snRNA digestion on spliceosomal complex formation (Section 3.2) it was noticed that digestion of the 5' end of U1 snRNA led to an unusually rapid accumulation of a low level of mRNA, to the extent that it was observed at the '0' time point. The '0' time points were typically taken 30s-120s after the addition of RNA to the splicing reaction (Figures 3.3C and 3.6C). To assess whether this was an artefact of RNase H digestion, β -globin RNA was incubated with RNase H and anti U1 oligonucleotide in the absence of NE (Figure 3.11B No Nuclear Extract sample). This showed that the exogenously-added RNase H did not digest the RNA to produce a band that was the same size as the mRNA. The '0' time point band was then excised from a gel (Figure 3.11B), amplified by PCR using the P16 and P17 oligonucleotides (Table 2.1), sequenced and found to contain the β -globin mRNA.

The observed effect was dependent on incubation with the Anti U1 oligonucleotide as a random 15nt oligonucleotide (Table 3.2) did not have the same effect when added to the same concentration (8 pmol/ μ l) as the Anti U1 oligonucleotide (Figure 3.11C). To see whether this effect was specific to the β -globin RNA the Ad1MW RNA was spliced in the same extract. The Ad1MW RNA is a modified version of the Ad1WW RNA wherein the upstream 5' SS has been mutated to leave an pre-mRNA with only one 5' SS (O'Mullane and Eperon, 1998). The effect of U1 knockdown was not observed on the Ad1MW substrate showing it to be substrate-specific (Figure 3.11D).

It was possible that the observed effect was due to induction of U1 snRNP-independent splicing (Crispino et al., 1994; Fukumura et al., 2009). When the 5' terminus of U1 snRNA is removed, a U1 snRNP-independent pathway of splicing may be able to occur in which splicing can occur very quickly on a small percentage of the pre-mRNA. If the effect of DNA oligonucleotide-mediated RNase H cleavage is due to U1-independent splicing, then the 2'O-methyl-modified α -U1 oligonucleotide that sequesters U1 snRNA should have the same effect. The α -U1 oligonucleotide inhibited splicing in a concentration-dependent manner but did not alter the kinetics of the reaction. In contrast Anti-U1 oligonucleotide-mediated RNase H digestion of the U1 snRNA inhibited splicing at the lower concentrations but then appeared to alter the kinetics of product formation at the higher concentrations. The efficiency of U1 snRNA RNase H digestion was assessed by running total nucleic acids purified from NE on a denaturing urea 6 % polyacrylamide gel and post staining with ethidium bromide. This showed that at the highest concentrations of Anti-U1 oligonucleotide full length U1

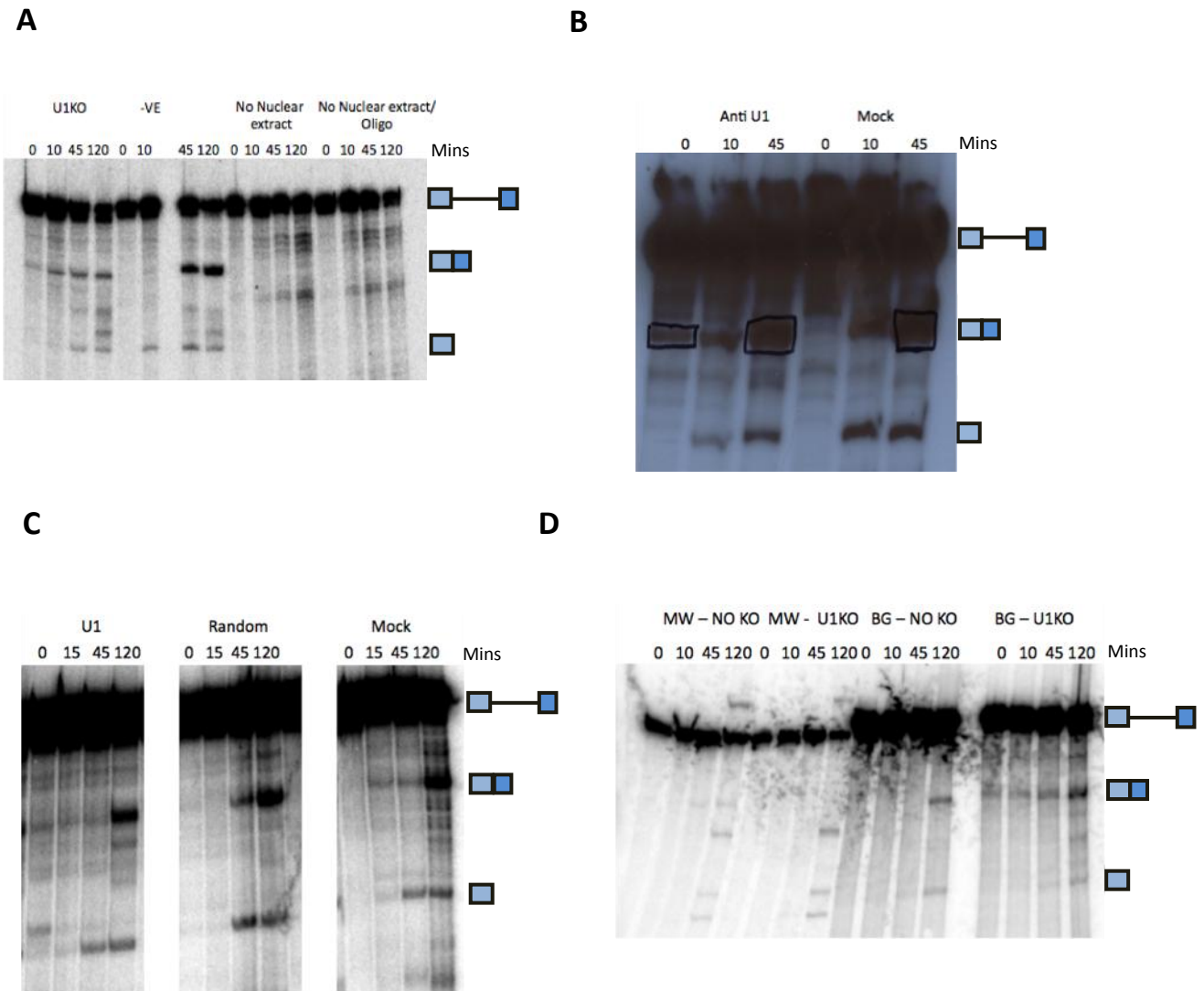
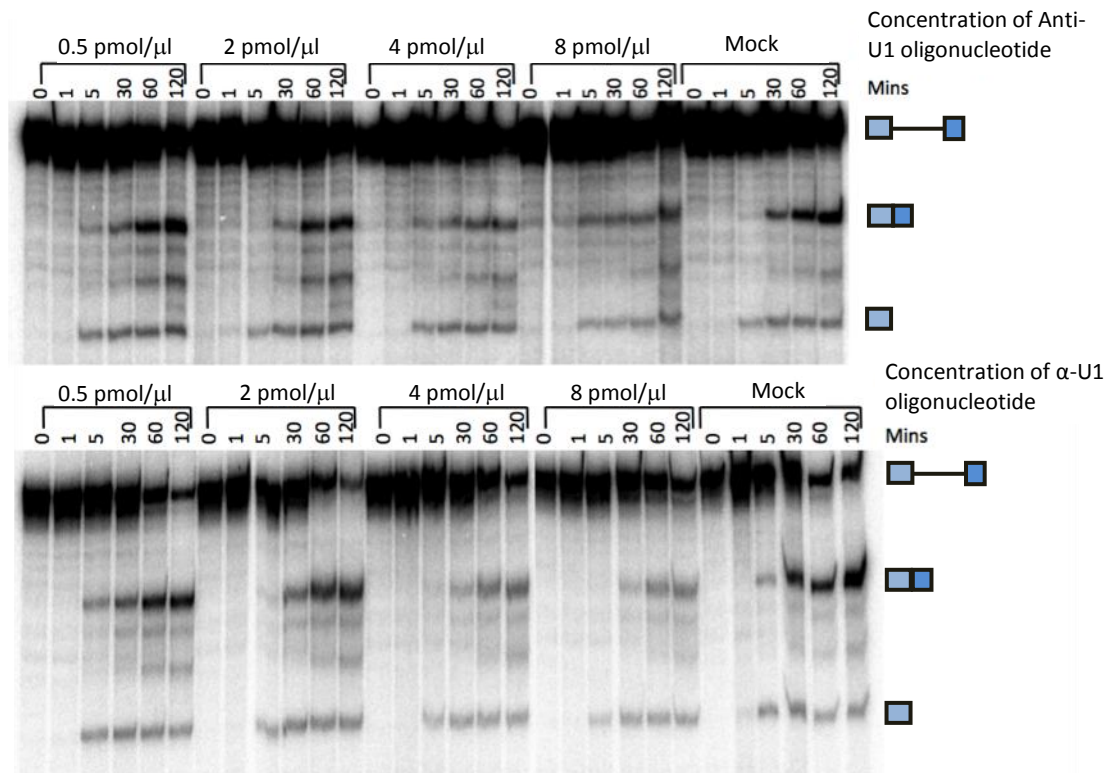


Figure 3.11: Detection of unusual products when U1 snRNA was degraded using RNase H. (A) Denaturing urea 6 % polyacrylamide gel showing that the effect of U1 snRNA degradation (U1 KO) cannot be reconstituted in the absence of NE (No Nuclear extract) or in the absence of NE and oligonucleotide (No Nuclear extract/ Oligo). (B) Denaturing Urea 6 % polyacrylamide gel indicating the bands that were excised and sequenced. (C) A denaturing urea 6 % polyacrylamide gel showing the splicing products produced by extracts digested with the Anti-U1 oligonucleotide (U1), a random oligonucleotide (Random) or no oligonucleotide (Mock). (D) A denaturing urea 6 % polyacrylamide gel showing the effect of U1 snRNA knockdown (U1KO) or no treatment (NO KO) on the splicing of the Ad1-MW and β -globin pre-mRNAs.



Anti-U1 oligonucleotide A-U1 oligonucleotide

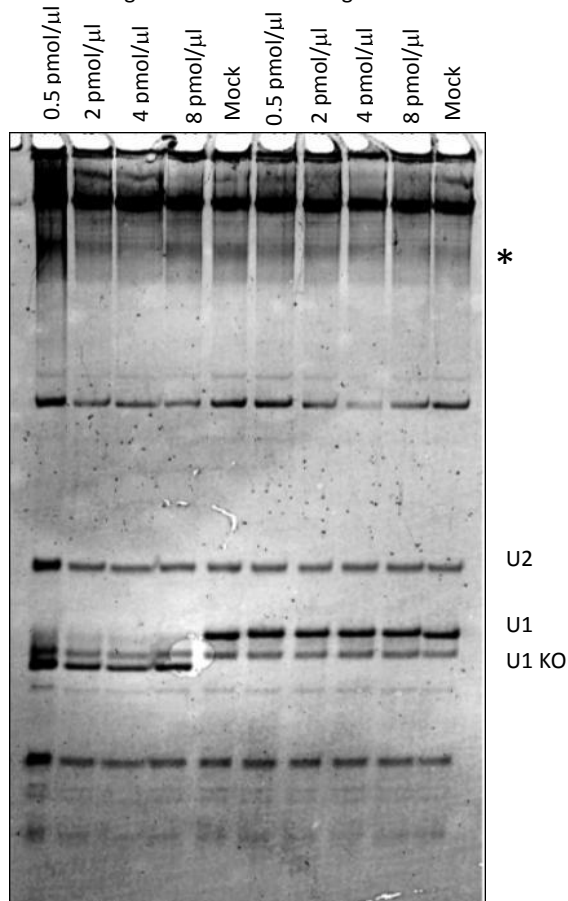


Figure 3.12: A comparison between the effects of knockdown and sequestration of U1 snRNA on splicing efficiency. The top panel shows time courses of a titration of the Anti-U1 oligonucleotide needed for RNase H-mediated degradation of U1 snRNA. The middle panel shows time courses of a titration of the α-U1 Oligonucleotide that sequesters the 5' end of U1 snRNA; gels in these two panels were denaturing urea 6 % polyacrylamide gels. The bottom panel shows an ethidium-stained denaturing urea 6 % polyacrylamide gel showing the snRNAs present in the NE used for splicing in the top two panels. The star indicates the possible smear of DNA.

snRNA was significantly reduced (Figure 3.12). The fact that U1 snRNP sequestration snRNA did not show the same effect as the experiments with DNA oligonucleotide-mediated RNase H cleavage clearly ruled out the possibility that the band was due to U1 snRNP-independent splicing.

Instead of U1 snRNP-independent splicing, the phenomenon might be explained if the U1 snRNA in the extract had the heterogeneous 5' ends that can be generated by RNase H cleavage (I. C. Eperon unpublished data) and have been detected in NE (Lerner et al., 1980). A set of snRNA molecules that have not been completely digested could bind to the β -globin RNA with lower affinity and dissociate readily during the rearrangements later in spliceosome formation. This hypothesis was supported by experiments showing that a U1 snRNP lacking its 5' terminus is able to bind an RNA specifically at the 5' SS (Rossi et al., 1996; Lund and Kjems, 2002; Du and Rosbash, 2002). When analysing the efficiency of snRNA digestion, a band was observed that migrated below full-length U1 snRNA but above the RNase H-cut U1 snRNA (Figure 3.12). It was hypothesised that this was a U1 snRNA which lacked part of its 5' terminus because of the extraction procedure and that it may be able to splice a subset of pre-mRNAs in the absence of full-length U1 snRNA. To determine whether this band was a short version of U1 snRNA, nucleic acids from NE were purified and digested with RNase H and oligonucleotides designed to be complementary to other regions of the U1 snRNA (Table 3.2). Figure 3.13A shows that the oligonucleotides that digested full-length U1 snRNA did not digest the unknown band (labelled with a question mark), thus showing this band was not a partially-digested U1 snRNA.

To understand whether the effect of U1 snRNA digestion on splicing was due to the RNase H-cut U1 snRNP having a lower affinity for the β -globin RNA and therefore being released more easily, a temperature gradient experiment was performed. The rationale behind this was that if the shortening of U1 snRNA led to an easier release from the spliceosome, then splicing may be able to occur at lower temperatures. Products were formed in the U1 snRNA-digested extract at temperatures as low as 1 °C (Figure 3.13B), suggesting that by removing the 5' end of U1 snRNA a major kinetic block in splicing had been removed. However, in the temperature gradient experiments, the 5' exon intermediate was not observed, unless the reaction was at a temperature permissive for splicing (Figure 3.13B). This meant that either the intermediate phase was very short-lived or that the products were not splicing-related.

To test whether the products observed after treatment of NE with Anti U1 oligonucleotide and RNase H were produced by splicing, a shortened transcript was produced that lacked the 3' exon and 3' SS AG dinucleotide. This transcript should only be able to go through the first transesterification reaction of splicing as there is no 3' SS AG for the second transesterification. This RNA was produced by transcription from a template of β -globin that had been digested using the SfcI restriction endonuclease which cut just upstream of the 3' SS AG. If the unusual products produced by U1 snRNA digestion were splicing products, then this RNA should produce splicing intermediates when incubated at 10 °C in U1 snRNA-digested NE. Figure 3.14A shows that the β -globinSfc RNA appeared to produce the 5' exon and intron lariat intermediates suggesting that the products were produced by splicing.

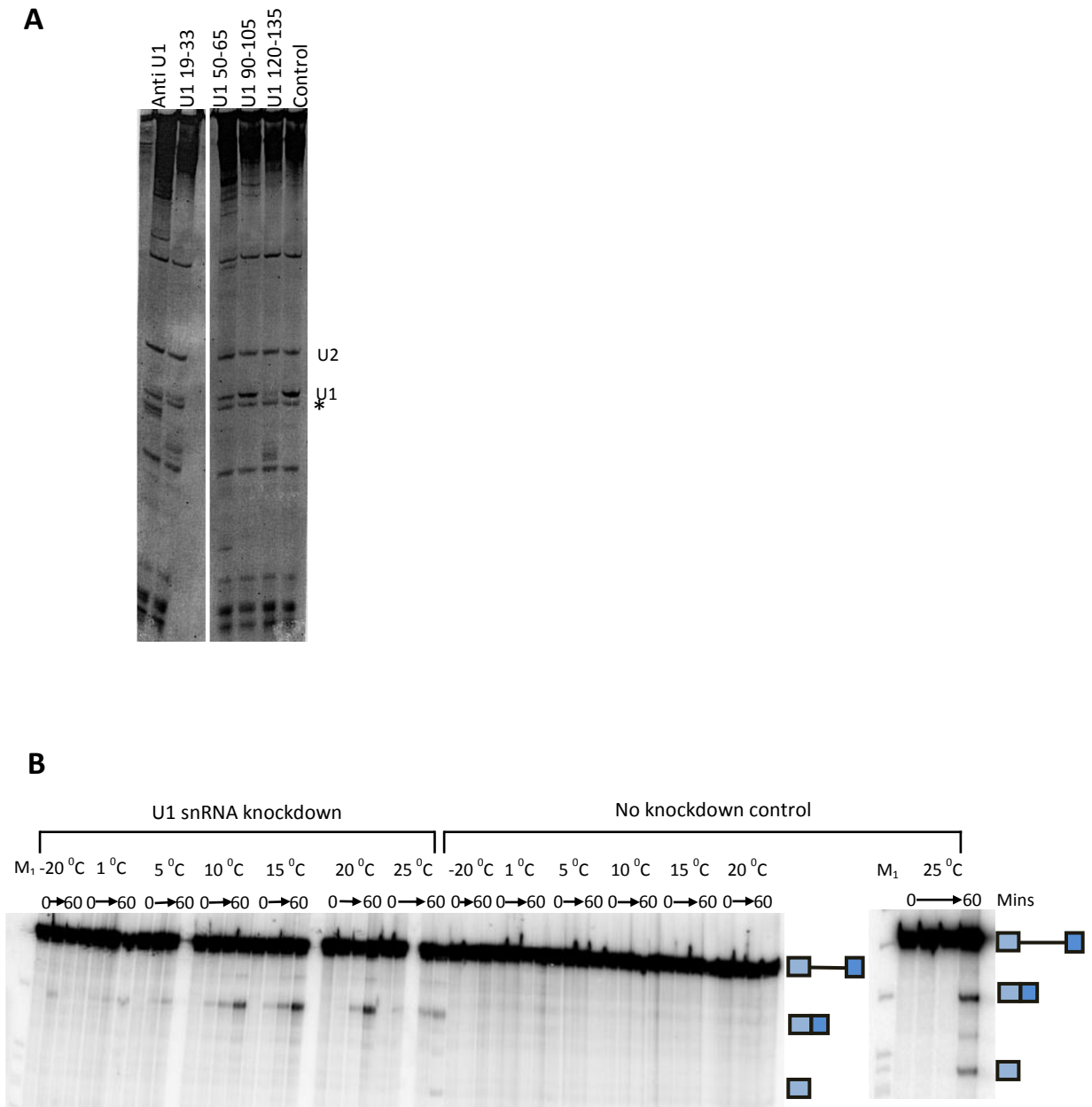


Figure 3.13: Characterization of the U1 snRNA knockdown-dependent products. (A) Ethidium-stained denaturing urea 6 % polyacrylamide gel showing RNase H-mediated digestion of U1 snRNA using oligonucleotides designed to cut at different points along the U1 snRNA. The star indicates the possible U1-related RNA **(B)** Products from splicing reactions carried out at different temperatures using either NE in which U1 snRNA has been digested with the Anti-U1 oligonucleotide and RNase H (U1snRNA knockdown extract) or NE that was without knockdown (no knockdown extract), time points of 0, 5, 15 and 60 minutes were taken.

To confirm that the lower band produced by β -globin in U1 snRNA-digested NE was a lariat, products of β -globin splicing in U1 snRNA-digested NE were run on a denaturing urea 12 % polyacrylamide gel alongside samples produced in non-digested NE.

However, this showed that the product thought to be mRNA was not the same size as the mRNA produced by splicing carried out at 30 °C (Figure 3.14B). The migration of the smaller fragment did not change, showing it cannot be a lariat as the intron-lariat produced by splicing at 30 °C clearly had an altered pattern of migration. This finding led to the conclusion that the products observed are in fact due to RNase H digestion of the pre-mRNA mediated by the Anti U1 oligonucleotide.

To confirm this, RNA from splicing reactions carried out using U1-digested and untreated NE at 10 °C and 30 °C were digested using the Pos-1 and Pos-2 DNA oligonucleotides (Table 3.2). The positions at which the two oligonucleotides were designed to cut are shown in the diagram in Figure 3.14C. The results of this experiment show that the Pos-1 oligonucleotide digested the pre-mRNA as expected but also digested the smaller RNA fragment (Labelled 1) produced in a splicing reaction carried out at 10 °C using extract where U1 snRNA had been digested (U1KO 10 °C sample). This showed that the band previously thought to be an intron-lariat contains the 5' end of the β -globin RNA and could therefore not be a lariat. The Pos-2 oligonucleotide stimulated cleavage of the pre-mRNA and the larger fragment (Labelled 2); as this oligonucleotide cuts in the intron of the β -globin pre-mRNA. this band cannot be mRNA as mRNA should not contain the intron.

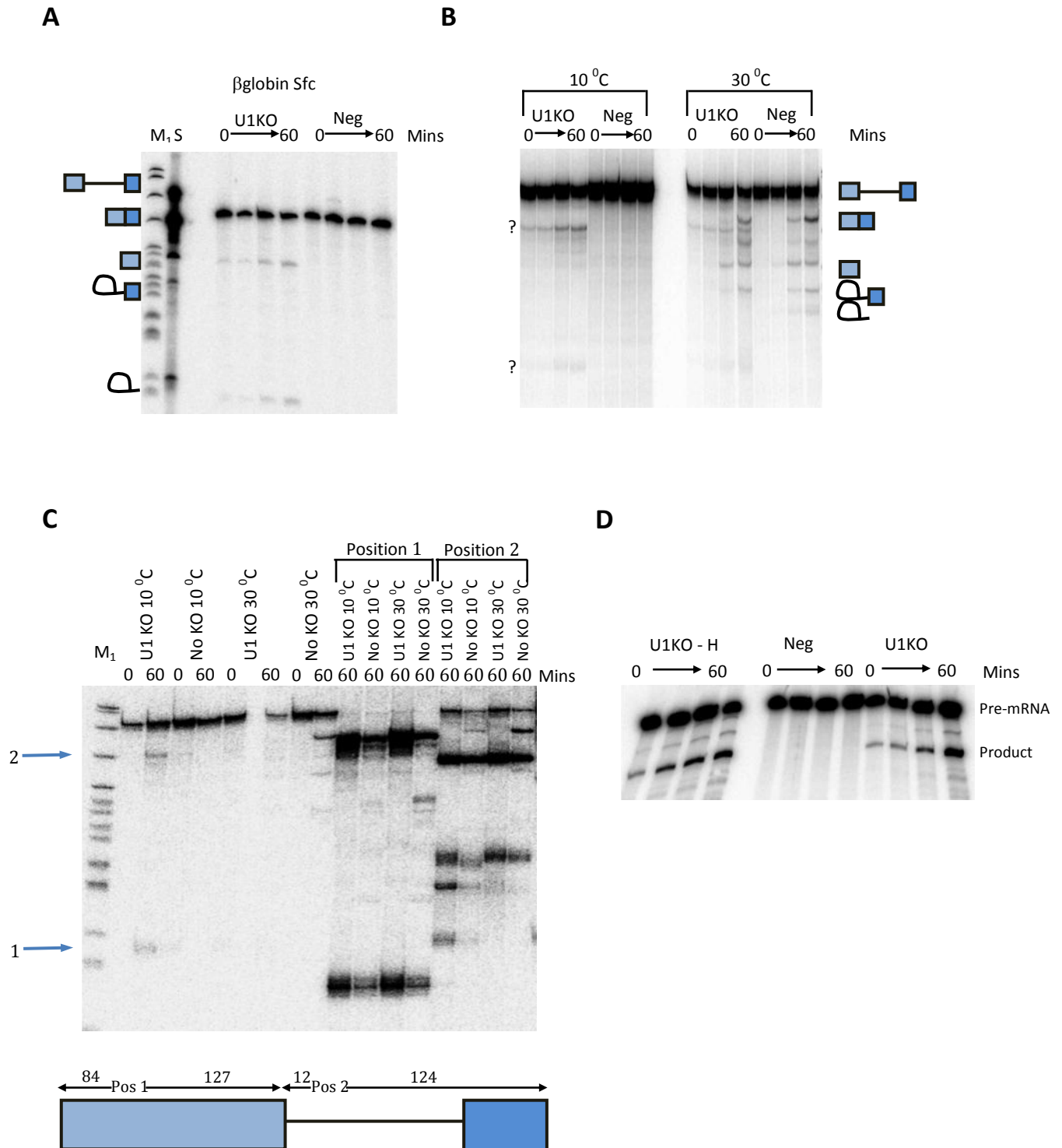


Figure 3.14: The products formed in a U1 knockdown reaction are not related to splicing. (A) Products formed by the β-globin-Sfc RNA when spliced in U1 snRNA knockdown or control extract at 10 °C, lane labelled S is a splicing reaction carried out using the β-globin RNA, used as a marker for the products of splicing. (B) Products formed by the β-globin RNA when spliced in U1 snRNA knockdown or control extract at both 10 °C and 30 °C run on a 12 % denaturing urea polyacrylamide gel. Time points of 0, 5, 15 and 60 minutes were taken. (C) Analysis of the U1 snRNA knockdown dependent products by RNase H mediated digestion of the β-globin RNA, panel one shows the results of the RNase H digestion run on a 12 % denaturing urea polyacrylamide gel whilst panel two illustrates the relative positions of the oligonucleotides used for digestion. (D) Products formed by the β-globin RNA after incubation at 10 °C in U1 snRNA knockdown extract (U1KO), control extract (Neg) or an extract treated with the oligonucleotide but not RNase H (U1KO-H). Time points of 0, 15, 30 and 60 minutes were taken and the samples were resolved on a denaturing urea 6% polyacrylamide gel.

In contrast, the mRNA produced by splicing of the β -globin pre-mRNA in the untreated extract incubated at 30 °C (No KO 30 °C), was left uncut by the Pos-2 oligonucleotide, but appeared to be cut by the Pos-1 oligonucleotide as would be expected for mRNA.

Judging by the size of the fragment labelled 1 the pre-mRNA of β globin was being cut at approximately nucleotide 115 relative to the 5' terminus when incubated in U1 snRNA digested extract. Interestingly, analysis of the pre-mRNA showed no obvious complementarity between this region and the Anti U1 oligonucleotide.

Addition of the oligonucleotide in the absence of exogenous RNase H yielded the same digestion products (Figure 3.14D). This, in conjunction with the results shown in Figure 3.11A, and the observation that addition of RNase H to the reaction did not increase the proportion of cleaved RNA (Figure 3.14D), suggested that the RNase H present within the NE was catalysing the digestion. The endogenous RNase H present in the NE thus appeared to have a lower requirement for complementarity than the added RNase H. Differences in the requirements of RNase H from different sources, for complementarity, have been observed before (for example Donis-Keller, 1979).

The results of sequencing the band excised in Figure 3.11B can be explained by the assay used. The RNA fragment cut out of the gel was reverse transcribed and then amplified using PCR. This PCR was done using primers complementary to the 5' and 3' terminus of the construct (P16 and P17) and so it selected for any low level of mRNA which might have been present.

3.5 The effect of DNase on splicing

Whilst investigating the complexes formed by mEGFP-U1-A in NE, preliminary experiments indicated that treatment of NE at 30 °C for 30 minutes in the presence of 1U/μl DNase (Promega) inhibited splicing completely. Further experiments showed that this inhibition was not due to RNase contamination, as the snRNAs were still present in NE treated with DNase (Figure 3.15A). Instead, it appeared the DNase was removing a large fragment of DNA, when run on a 10 cm long denaturing urea 6 % polyacrylamide gel, indicated by a blue arrow (Figure 3.15A).

Experiments showed that this effect was dependent on the activity of DNase, inactivation of DNase by incubation at 80 °C for 5 minutes prior to addition to NE led to the loss of any effect. Addition of the buffer that the DNase was stored in (10mM Hepes pH 7.5, 50 % v/v glycerol, 10mM CaCl₂, 10mM MgCl₂) also showed no effect (Figure 3.15B). In contrast, active DNase abolished splicing and EcoR1 appeared to inhibit splicing (Figure 3.15B). When the nucleic acids present in the NE were analysed, it appeared that the DNase had removed the large band whilst the EcoR1 had reduced its intensity in comparison to the negative control.

The nucleic acid band that was removed by DNase treatment appeared to be heterogeneous since when total nucleic acids from NE were run on a longer 20 cm denaturing urea 6 % polyacrylamide gel the band was no longer visible, instead there appeared to be a smear (Figure 3.12, indicated by blue arrow). This has led to the hypothesis that DNA fragments are being produced by the extraction procedure, probably by shearing of the genomic DNA.

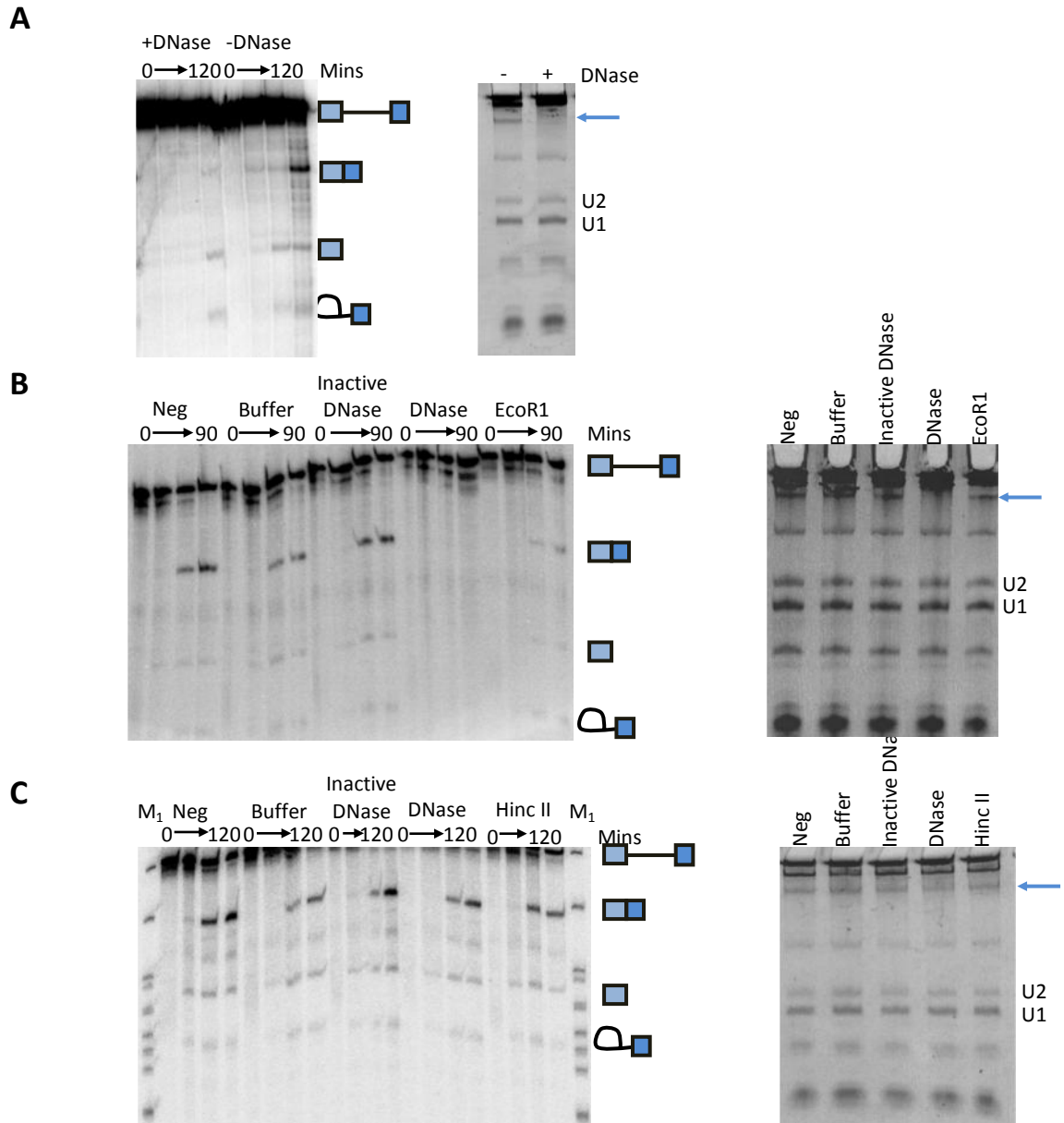


Figure 3.15: The DNase effect. **(A)** The First Panel shows a denaturing urea 6 % polyacrylamide gel of the splicing products produced in the presence (+DNase) and absence of DNase (-DNase; -DNase sample is same as Mock sample in Figure 3.11C). Time points of 0, 15, 45 and 120 minutes were taken. The adjacent panel shows an ethidium-stained denaturing urea 6 % polyacrylamide gel of the nucleic acids purified from the NE used for splicing in the first panel. **(B)** Inactivation of DNase stops inhibition. The first panel shows the splicing products formed at 0, 15, 45 and 90 minutes run on a denaturing urea 6 % polyacrylamide gel. The NE was either untreated (Neg), treated with DNase buffer (Buffer), inactive DNase (Inactive DNase), DNase (DNase) or EcoR1 (EcoR1). The adjacent panel shows an ethidium-stained denaturing urea 6 % polyacrylamide gel; it shows the effect of the treatments in the first panel on the nucleic acids in the NE. **(C)** Failure of DNase to inactivate splicing. The First panel shows a denaturing urea 6 % polyacrylamide gel of the splicing products under different conditions as in **B** with the addition of HincII treatment. Timepoints of 0, 15, 45 and 120 minutes were taken. The adjacent panel shows an ethidium-stained denaturing urea 6 % polyacrylamide of the nucleic acids purified from the NE used for splicing in the first panel.

This line of research was abandoned, as results were not always consistent. On some occasions DNase treatment had no effect on splicing efficiency (Figure 3.15C) and analysis of the nucleic acids in the NE after DNase treatment showed the DNase had not removed the large fragment. At the time it was noted that many undergraduate students were also using the DNase, and previous batches had been inactivated through incorrect use.

Summary

Through the work discussed in this chapter it has been possible to define the spliceosomal complexes formed on pre-mRNAs derived from the β -globin pre-mRNA. Interestingly this is the first time that the B^{ACT} complex has been detected on a full-length pre-mRNA by native LMP-agarose gel electrophoresis. The effect of addition of phosphatase inhibitor on spliceosomal complex formation and the effect of DNase on splicing efficiency have also been investigated. An artefactual effect on the apparent splicing of β -globin caused by the RNase H mediated digestion of U1 snRNA was also investigated.

Chapter 4. Investigations into the number of U1 snRNPs bound to pre-mRNAs containing multiple 5' SS

4.1 Assays to determine whether fluorescent fusion proteins are functional

4.2 Construction, characterization and labelling of the pre-mRNAs

4.3 Preparation of samples for TIRF microscopy

4.4 One U1 snRNP associates with pre-mRNAs containing a single 5' SS under E complex conditions

4.5 Two U1 snRNPs probably associate with pre-mRNAs containing two 5' SS under E complex conditions

4.6 Loss of one U1 snRNP from pre-mRNAs containing two 5' SS by A complex

4.7 Loss of U1 snRNP is dependent on the presence of a 3' SS downstream and is inhibited by the presence of a 3' SS between the two 5' SS

Introduction

The objective of the work described in this chapter was to answer one of the questions posed in the Introduction; namely, how many U1 snRNPs bind pre-mRNAs in which 5' SS selection depends on the relative position of the sites. The number of U1 snRNPs bound to individual RNAs was assessed using TIRF microscopy.

4.1 Assays to determine whether fluorescent fusion proteins are functional

To achieve this goal, a method to label the U1 snRNP was required. To do this, nuclear extracts were made from HeLa cells transfected with fluorescent fusion protein constructs. The fluorescent fusion protein constructs, mEGFP-U1-A and mEGFP-U1-C had recently been created within the laboratory by Dr L.P. Eperon. They were created by fusing the entire open reading frame of U1-A or U1-C to the C-terminus of mEGFP, which carries a mutation (A206K), known to reduce homodimerisation of EGFP (Zacharias et al., 2002). As part of these constructs, a linker encoding (glycine)₃serine was introduced between the C terminus of mEGFP and N terminus of the protein of interest.

The U1 snRNP specific factor, U1-A was used as a marker for the U1 snRNP because of its position within the U1 snRNP; it is bound to stem loop II of the U1 snRNA, and therefore is not close to the region of U1 snRNP that binds the pre-mRNA (Weber et al., 2010). U1-A is neither needed for U1 snRNP's association with the 5' SS (Heinrichs et al., 1990), nor for spliceosomal complex formation in NE (Will et al., 1996). Therefore it was assumed that tagging of U1-A with mEGFP would not interfere with the ability of U1 snRNP to bind the 5' SS.

In contrast, the protein U1-C, as discussed in Section 1.3, is very important for 5' SS selection as it is able to select a 5' SS in the absence of the 5' terminus of U1 snRNA. It was possible that addition of a mEGFP-tag would interfere with its function but, due to its role in 5' SS selection, it was thought interesting and was pursued.

To test whether the fusion proteins were part of the U1 snRNP, nuclear extracts were made from HeLa cells expressing either mEGFP-U1-A or mEGFP-U1-C and then biotin affinity precipitations were performed using the α -U1 oligonucleotide (Table 2.1), as described in section 2.2.6. It was expected that if the fluorescent fusion proteins were part of the U1 snRNP then precipitation of the U1 snRNA using this biotinylated oligonucleotide should also precipitate the fluorescent fusion proteins.

When pull-downs were performed using a NE containing mEGFP-U1-A it was clear that mEGFP-U1-A had been brought down as part of the U1 snRNP; detection of U1-A by Western blot using the α -U1-A(a) antibody (Table 2.3) showed that both the endogenous U1-A and mEGFP-U1-A appeared in the pellet. The mEGFP-U1-A band was identified due to its size and its absence when commercial extract was used for the biotin affinity precipitation. Affinity precipitation of both bands was dependent upon the addition of the α -U1 oligonucleotide to the reaction, showing that the signal was not due to non-specific absorption onto the neutravidin beads (Figure 4.1A). The specificity of the pull-down was confirmed by analysis of the snRNA that was precipitated by the procedure. When the RNAs precipitated by a similar pull-down were purified by treatment of the reaction with proteinase K, ethanol precipitated, then run on a denaturing urea 6 % polyacrylamide gel and silver-stained, only a single band was detected. The band was identified as U1 snRNA by an affinity precipitation

performed using NE in which the 5' terminus of U1 snRNA had been digested using RNaseH and the Anti-U1 oligonucleotide (Table 3.2) at a concentration of 8 pmol/ μ l as described in Section 2.1.6. This treatment led to the loss of the band detected by silver-staining (Figure 4.1B).

In contrast, when biotin affinity precipitations were performed on NE prepared from HeLa cells expressing mEGFP-U1-C, mEGFP-U1-C could not be detected in the precipitate sample by Western blotting using the α EGFP antibody (Table 2.3; Figure 4.1C). The levels of endogenous U1-C could not be determined due to a lack of a suitable antibody; however latter experiments performed by others in the laboratory indicated that endogenous U1-C can be precipitated using this oligonucleotide when judged by silver-staining (Mrs R. Dumas, unpublished data). As a control for the biotin affinity precipitation, a second Western blot was performed on the pull-down using the α -U1-A(a) antibody; this showed efficient precipitation of U1-A, confirming that U1 snRNP had been precipitated, but that mEGFP-U1-C was unable to associate with it (Figure 4.1C).

To determine whether mEGFP-U1-A containing U1 snRNPs were able to bind pre-mRNA, biotin affinity precipitations were performed using biotinylated globin C pre-mRNA. These experiments clearly showed that this RNA precipitated both endogenous U1-A and mEGFP-U1-A. Biotin pull-downs performed in NE in which the 5' terminus of U1 snRNA had been digested, as above, showed that this precipitation was dependent on the ability of U1 snRNA to anneal to the 5' SS (Figure 4.1D). It was noted that the difference in the relative amounts of mEGFP-U1-A and endogenous U1-A in Figures 4.1A and 4.1D were due to the use of different extracts.

Immunoprecipitations of radiolabelled globin C pre-mRNA were performed using the α -EGFP antibody (Table 3.2) and splicing reactions containing NE-expressing fluorescent fusion proteins. Any of the fluorescent proteins that were able to associate with pre-mRNA were expected to precipitate the radiolabelled globin C pre-mRNA. Immunoprecipitations from NE-expressing mEGFP- U1-A, which had been pre-depleted of ATP by incubation at 30 °C for 30 minutes, precipitated far more globin C pre-mRNA than NE expressing EGFP or NE not expressing any fluorescent fusion protein (Figure 4.1E). This showed mEGFP-U1-A-containing snRNPs are able to associate with pre-mRNA and thus precipitate it. In contrast, when an immunoprecipitation was performed on a splicing reaction made using mEGFP- U1-C containing NE, which had been depleted of ATP, no precipitation of globin C pre-mRNA was observed. The positive control for the immunoprecipitation, mEGFP-SF2 (SFRS-1), did precipitate globin C pre-mRNA, showing that the immunoprecipitation worked (Figure 4.1F), but that mEGFP- U1-C was not active in the assay.

Overall, this data shows that mEGFP-U1-A is incorporated into the U1 snRNP and that U1 snRNPs containing mEGFP-U1-A are capable of binding pre-mRNA, but that mEGFP-U1-C is not incorporated into the U1 snRNP and is not functional. Due to these findings, only mEGFP-U1-A was used for further experiments.

4.2 Construction, characterization and labelling of the pre-mRNAs

The pre-mRNAs used in this study were derived from the β -globin or Ad1 series of constructs (Figure 4.2; Eperon et al., 1993; O'Mullane and Eperon, 1998). As well as

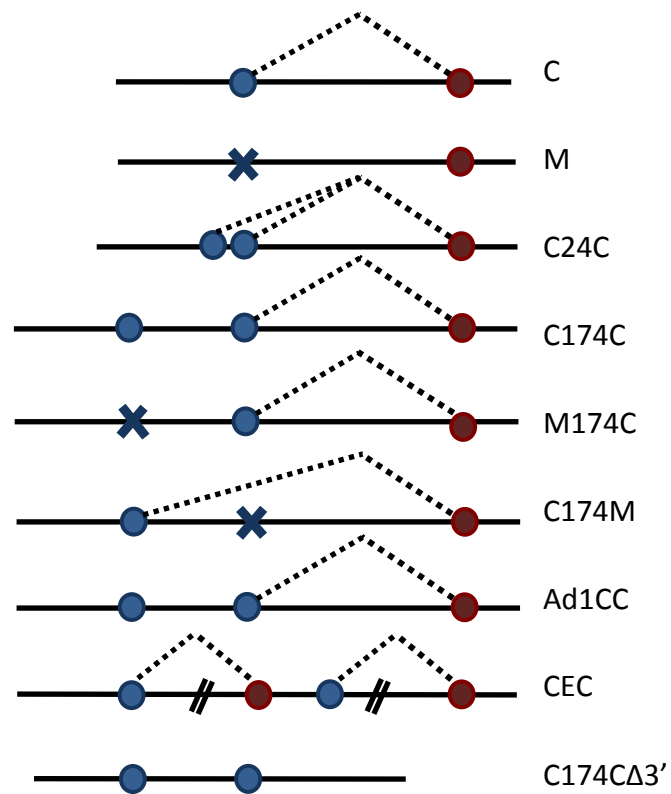


Figure 4.2: Representation of pre-mRNAs. Blue spots indicate 5' SS, red spots indicate 3' SS, blue crosses indicate mutated 5'ss and dashed lines indicate the observed pattern of splicing.

the globin C and globin M pre-mRNAs described in Section 3.2, other pre-mRNAs with more than one 5' SS were used. Previously in the laboratory, a series of RNAs containing multiple 5' SS had been constructed based upon the rabbit β -globin construct which contained exon 2, the full length intron 2 and exon 3. These pre-mRNAs had 5' splice sites separated by different numbers of nucleotides and they contained either the wild-type rabbit β -globin exon two 5' SS, or a consensus 5' SS (Eperon et al., 1993). The pre-mRNAs used from this set were the C24C, C174C and C175G pre-mRNAs, where C stands for the consensus 5' SS, G stands for wild-type β -globin 5' SS and the number indicates the number of nucleotides separating the two 5' SS.

As the globin C and globin M pre-mRNAs had been made by site-directed mutagenesis of the shortened version of β -globin pre-mRNA (Skordis et al., 2003), the intron of the C24C, C174C and C175G constructs was also shortened to make them the same length. This had the added bonus of making the pre-mRNAs easier to work with. To shorten the intron, two short fragments were amplified and then joined by PCR. The first fragment spanned from the 5' end of the construct to just downstream of the second 5' SS and was amplified using the P16 (Table 2.1) and β g ivs topo OS (Table 4.1) oligonucleotides. The second fragment was produced from the shortened version of the β -globin intron and stretched from the 5' end of the intron to the 3' end of the construct and was amplified using the β g ivs topo R (Table 4.1) and β g E3 no 5' SS (Table 2.1) oligonucleotides. These two fragments were extracted from an agarose gel and then added to a PCR reaction together. The two fragments were designed to have

a region of 47nt that overlapped, allowing them to anneal; P16 and β G E3 no 5' SS were then used to amplify the shortened versions of C24C, C174C and C175G from the two fragments. These PCR reactions were carried out using the Phusion-Hot Start DNA polymerase (Finnzymes) due to its proof-reading ability. Once amplified, the fragments were cloned into the PCR-vector using a Topo-kit (Invitrogen) as per manufacturer's instructions and sequenced.

Name	Sequence
Bg ivs topo OS	ATAACATGAATTTTACAATAGCG
Bg ivs topo R	ACCCTTGATTGTTCTTTCTTTGCTA
C174MF	ACCAGAAACGCGCCGAACCTTTGGGGACCCTTGATTGTTC
C174MR	GAACAATCAAGGGTCCCCAAAAGTTGCGCGCGTTTCTGGT
M174CF	ACAAGCTGCACGTGGATCCCCGATCCTTCGGTACACCACG
M174CR	CGTGGTGTACCGAAGGATCGGGGATCCACGTGCAGCTTGT
BGE3/E2repeat	CCTGGGCAACGTGCTGGTTAGGACCCACAGGTTCTTCGAG
BGE3 3' Rep	GATCCAAGAGTACTAGTTAGACCGCGAACTTACCTGCCAAAATG

Table 4.1: Sequences of oligonucleotides. Oligonucleotides used for construction of transcripts and site directed mutagenesis.

The mutants of the C174C, C174M and M174C were made by site-directed mutagenesis; M stands for a mutated 5' SS. To create the M174C pre-mRNA, the upstream 5' SS was removed using the M174CF and M174CR oligonucleotides, whilst mutagenesis to create C174M was performed using the C174MF and C174MR oligonucleotides (Table 4.1) using standard site-directed mutagenesis conditions (Section 2.4.4).

Amplification of the β -globin series of constructs by PCR for transcription was done with the P16 and β G E3 no 5' SS oligonucleotides (Table 2.1). To create a version of the C174C that lacked the 3' SS (C174C Δ 3'), the construct was amplified using the P16 and Glo C-3' end oligonucleotides (Table 2.1), the latter of which primed from just upstream of the BPS in the shortened β -globin intron.

As a control, to investigate whether any effects that were seen were substrate-specific, the Ad1-CC RNA (O'Mullane and Eperon, 1998) was also used, without modification. Polymerase chain reaction amplification of this construct for transcription was performed using the Ad1 Trx F and Ad1 Trx R oligonucleotides (Table 2.1).

As the length of the intron in the C24C and C174C constructs had been shortened it was important to ensure that the splicing pattern had not been altered. As expected, globin C spliced efficiently to the consensus 5' SS whilst globin M spliced with low efficiency, utilising a known weak cryptic 5' SS in the β -globin pre-mRNA. C24C spliced with low efficiency to both 5' SS, as expected from previous work (Eperon et al., 1993), despite both 5' SS being consensus sequences (Figure 4.3A). The efficiency of splicing had previously been shown to be enhanced by cleavage of the 5' terminus of U1 snRNA, thereby reducing the complementarity between U1 snRNA and the 5' SS. This led to the hypothesis that the observed inhibition of splicing was because two U1 snRNPs that are too close together interfere with each other and inhibit 5' SS use (Eperon et al., 1993).

In contrast, as expected, C174C spliced almost exclusively to the downstream 5' SS with only a small amount of the 5'exon intermediate from the upstream 5' SS being produced, clearly showing the position effect. The C174M and M174C variants of C174C spliced only to the 5' SS that remained showing that both 5' SS are intrinsically available for splicing. The Ad1 CC pre-mRNA also displayed a position effect, as the downstream 5' SS was heavily favoured with a low level of splicing occurring using the upstream 5' SS (Figure 4.3A).

Quantification of the efficiency of splicing observed for the C174C and M174C in this experiment yielded an interesting observation. The efficiency of splicing of the two pre-mRNAs was indistinguishable, meaning that the presence of an upstream 5' SS in this context has not enhanced splicing to the downstream 5' SS (Figure 4.3B). This appeared to contradict conclusions drawn from experiments using derivatives of the human β globin gene, in which the presence of an upstream 5' SS was found to enhance splicing to a downstream 5' SS (Hicks et al., 2010). This difference in results is not due to a difference in the distance between the two 5' SS, as in the constructs used by Hicks et al (2010) the two 5' SS were separated by 180nt. However, it may be due to the sequence used for the downstream 5' SS; Hicks et al (2010) used a 5' SS sequence that had two mis-matches to the U1 snRNA whilst the results in Figure 4.3B were produced using a 5' SS sequence with perfect complementarity. As a result, it cannot be ruled out that the lack of an effect is due to the strength of U1 snRNP binding.

To be able to observe the pre-mRNA using the TIRF microscope a labelling technique was needed. To this end, fluorescently-labelled oligonucleotides (Eurogentec) were used. To label pre-mRNAs derived from the β -globin construct, the Bg-5'-Cy5 oligonucleotide was used (Table 4.2). This oligonucleotide was designed to be complementary to the first sixteen nucleotides of the 5' exon and comprised a mixture of 2'O-methyl and LNA nucleotides. These two modifications were used as they are resistant to nuclease digestion and increase the stability of hybrids formed with

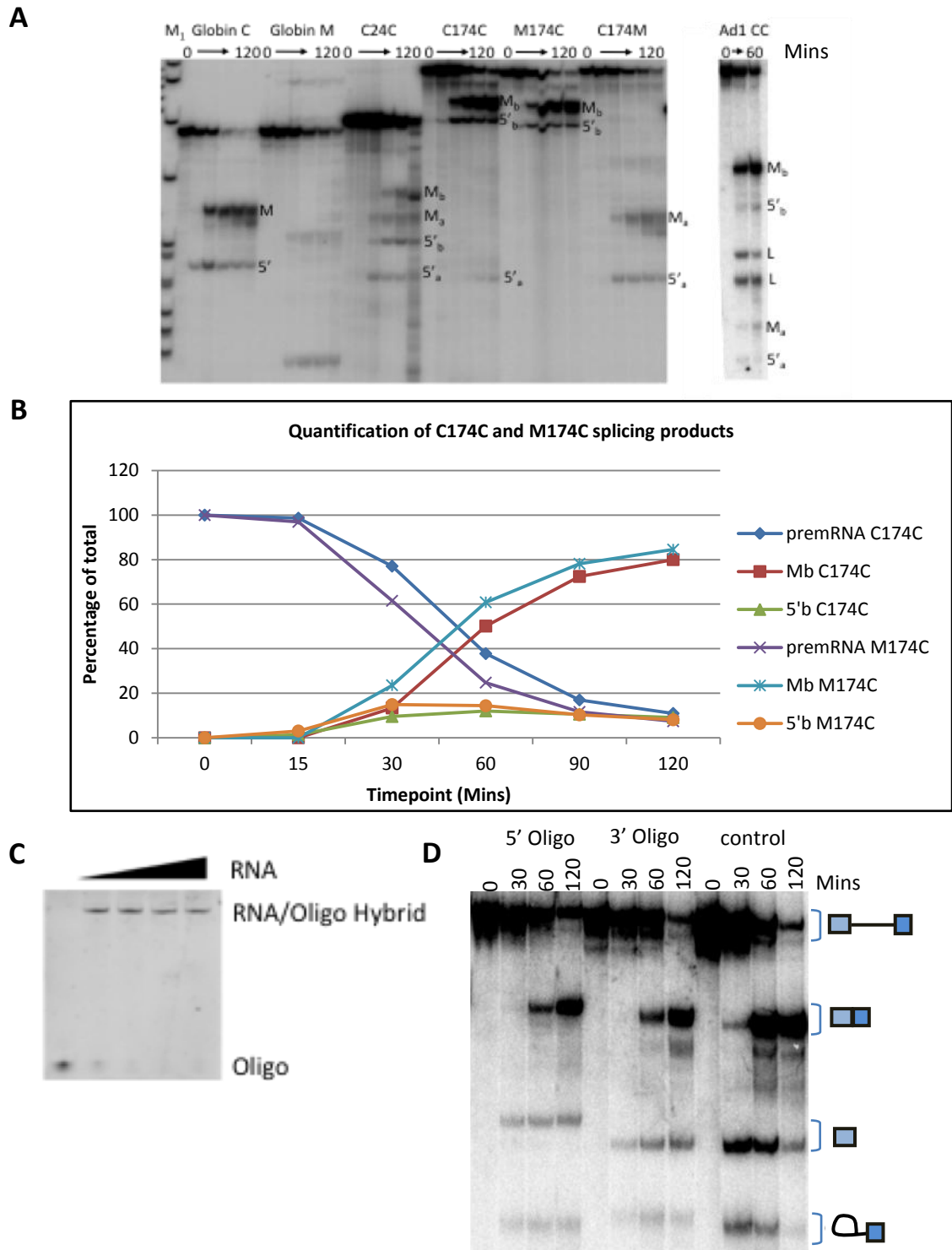


Figure 4.3: Labelling of the RNA. (A) Splicing of the pre-mRNAs indicated with time points of 0, 15, 30, 60, 90 and 120 minutes (globin series) or 0, 30 and 60 (Ad1 CC), 'M' stands for mRNA, '5'' stands for 5'exon and 'L' stands for lariat; when there are two possible 5' exons or mRNAs, those produced by the upstream 5' SS are labelled $M_a/5'_a$, whilst those produced by the downstream 5' SS are labelled $M_b/5'_b$. (B) Quantification of the C174C and M174C time courses shown in Figure 4.3A, corrected for the number of guanosine nucleotides in each product. (C) Hybridisation of fluorescent biotinylated oligonucleotide to pre-mRNA. The concentration of RNA ranged from 0.85 μ M-1.5 μ M (L-R). Samples were run on a denaturing urea 6 % polyacrylamide gel before being detected by excitation of the fluorophore using a Typhoon imager. (D) Splicing carried out with 50 nM radiolabelled pre-mRNA hybridised to either Bg-5'-Cy5 (5'Oligo) or Bg-3'-Alexa488 (3'Oligo); 'control' was spliced with a low level of radiolabelled RNA. Splicing products with complementarity to the oligonucleotides show attenuated migration.

complementary RNA (Cotten et al., 1991; Koshkin et al., 1998). To label the Ad1-CC RNA, Oligo 9 was used (Table 4.2), this oligonucleotide was designed to be complementary to the first fifteen nucleotides of the Ad1-CC pre-mRNA and was made of 2'O methyl-modified nucleotides.

Annealing of these oligonucleotides to complementary RNAs was stable enough to resist electrophoresis in 7M urea. This allowed assessment of the proportion of free oligonucleotide in an annealing reaction by denaturing urea 6 % polyacrylamide gel electrophoresis. This was useful, as it allowed analysis of the annealing efficiency without the interpretation being complicated by RNA secondary structures (Figure 4.3C); annealing reactions were carried out as described in Section 2.3.3.

Name	Sequence
Anti-U1-Cy5	Cy5- UGCCAGGUAAGUAU -biotin
Bg-5'-Cy5	Cy5- UAGACAACCAGCAGCC -biotin
Bg-3'-Alexa 488	Alexa488- ACCAAAATGAUGAGAC
Oligo 9	Cy5- ACCUGCAGGCAUGCA -biotin
Ad1-Intron	Cy3- TGCAGCAAGCTTGACAAC

Table 4.2: Sequences of fluorescent oligonucleotides. The 2'O methyl modified nucleotides are shown in bold lettering, the LNA modified nucleotides are shown in non-bold lettering, Cy5 denotes a Cyanine-5-Dye, Cy3 denotes a Cyanine-3-Dye, Alexa488 denotes an Alexa fluor 488 and biotin denotes a biotin group.

To test the effect that annealing the oligonucleotides to the pre-mRNA had on splicing, standard splicing reactions were carried out with 50 nM globin C pre-mRNA annealed to the Bg-5'-Cy5. To allow detection of the splicing products the annealing reaction was carried out in the presence of a small amount of radiolabelled globin C pre-mRNA, as well as the 50 nM cold globin C pre-mRNA. Analysis of the splicing products formed clearly shows that the level of splicing is only slightly inhibited when compared to the "no oligonucleotide" control, which was spliced without the addition of cold globin C pre-mRNA. Interestingly, the oligonucleotide did not appear to be dislodged from the

RNA as it was spliced as RNA products that were annealed to the oligonucleotide migrated more slowly in the gel. This was useful since it allowed confirmation that the oligonucleotides were annealing to the correct place on the RNA. The oligonucleotide, Bg-5'-Cy5, altered the migration of the pre-mRNA, mRNA and 5'exon but not the 3'exon-lariat. In contrast an oligonucleotide designed to anneal to the 3' exon, Bg-3'-Alexa 488 (Table 4.2), altered the migration of the pre-mRNA, mRNA and 3'exon-lariat but did not shift the 5'exon (Figure 4.3D).

As well as being fluorescently labelled at the 5' terminus, the Bg-5'-Cy5 and Oligo 9 oligonucleotides had a biotin group attached to the 3' terminus; this was important as it allowed the pre-mRNA to be recruited to the surface of the slide. As described in Section 1.8, TIRF microscopy relies upon an evanescent field that decays exponentially. This is important, as it leads to a loss of background signal from the fluorophores in the rest of the sample, but it also means that a method must be used to attach the complex to the surface. As described in Section 2.3.5, the surface of the slide had been modified to bind biotin through streptavidin. Therefore the biotin group on the oligonucleotide allowed deposition of the oligonucleotide-annealed pre-mRNA onto the surface of the slide. Any components of the NE that were associated with the pre-mRNA, such as the spliceosome, would therefore be brought to the surface of the slide.

4.3 Preparation of samples for TIRF microscopy

So that data obtained by TIRF microscopy was comparable, all experiments detailed here were carried out using a single NE prepared from HeLa cells transfected with the mEGFP-U1-A expression plasmid. Under standard conditions, this extract spliced

efficiently, when compared to commercial HeLa NE (Figure 4.4A). It also formed spliceosomal complexes when incubated with 50 nM globin C RNA and stalled at A complex using anacardic acid or the α -U6 oligonucleotide (Figure 4.4B).

The level of mEGFP-U1-A was assessed using a fluorescent Western blot (Figure 4.4C) as described in Section 2.2.5, using the α U1-A(b) antibody (Table 2.3). In this case, the antibody α U1-A(a) could not be used as it was not detectable using the α Mouse fluorescent secondary antibody (Table 2.3). This Western blot showed that the levels of endogenously expressed U1-A and exogenously expressed mEGFP-U1-A were approximately equal.

To set up a splicing reaction for analysis by TIRF microscopy, the mEGFP-U1-A-containing NE was either depleted of ATP by pre-incubation at 30 °C for 30 minutes in order to stall at E complex, or was pre-incubated in the presence of anacardic acid or the α -U6 oligonucleotide (Table 2.2) at 30 °C for 15 minutes to stall at A complex. This was then added to a splicing reaction containing 50 nM fluorescent oligonucleotide-labelled pre-mRNA and a splicing buffer containing the components detailed in Section 2.3.4.

The components of this reaction were different to those used in a standard splicing reaction, as they were the conditions used within the laboratory for assembly of splicing reactions for TIRF analysis. Preliminary experiments using standard splicing reaction conditions were attempted, but showed poor results; this was later found to be due to human error. To ensure the new conditions were not sub-optimal, splicing of globin C under standard splicing conditions and those used for single molecule TIRF

experiments were compared (Figure 4.4D). These results show that the conditions used for single molecule TIRF experiments were not detrimental to splicing efficiency.

Once assembled, the splicing reactions were incubated at 30 °C for 30 minutes to allow spliceosomal complex assembly, before being diluted in Dilution buffer (Section 2.5.1) so that the pre-mRNA concentration was approximately 3 pM. This sample was then ready for imaging with the TIRF microscope.

4.4 One U1 snRNP associates with pre-mRNAs containing a single 5' SS under E complex conditions

Once imaged, as described in Section 2.3.6, each field of data was analysed for co-localization. This entailed creating a “sum image” of the first 20-30 frames from the beginning of excitation for each channel using Image J and then overlaying the two images using Adobe Photoshop. The overlaying of the two images was done based on a known standard, because the beam splitter that was used to separate the two wavelengths of emission light relied upon mirrors; if one of these was not positioned correctly, then the area of the sample detected on the chip of the camera would be different for the two wavelengths.

The image standards used were fluorescent beads; these fluoresced in both channels upon excitation with the Argon Ion 488 nm laser (Table 2.4). This allowed a signal of fluorescence that came from the same spot to be collected in both channels. When overlaying, the distance one image had to be moved relative to the other to show co-localized spots was recorded. An example of a bead overlay experiment is given in Figure 4.5A. As long as the mirrors within the beam splitter were not altered between

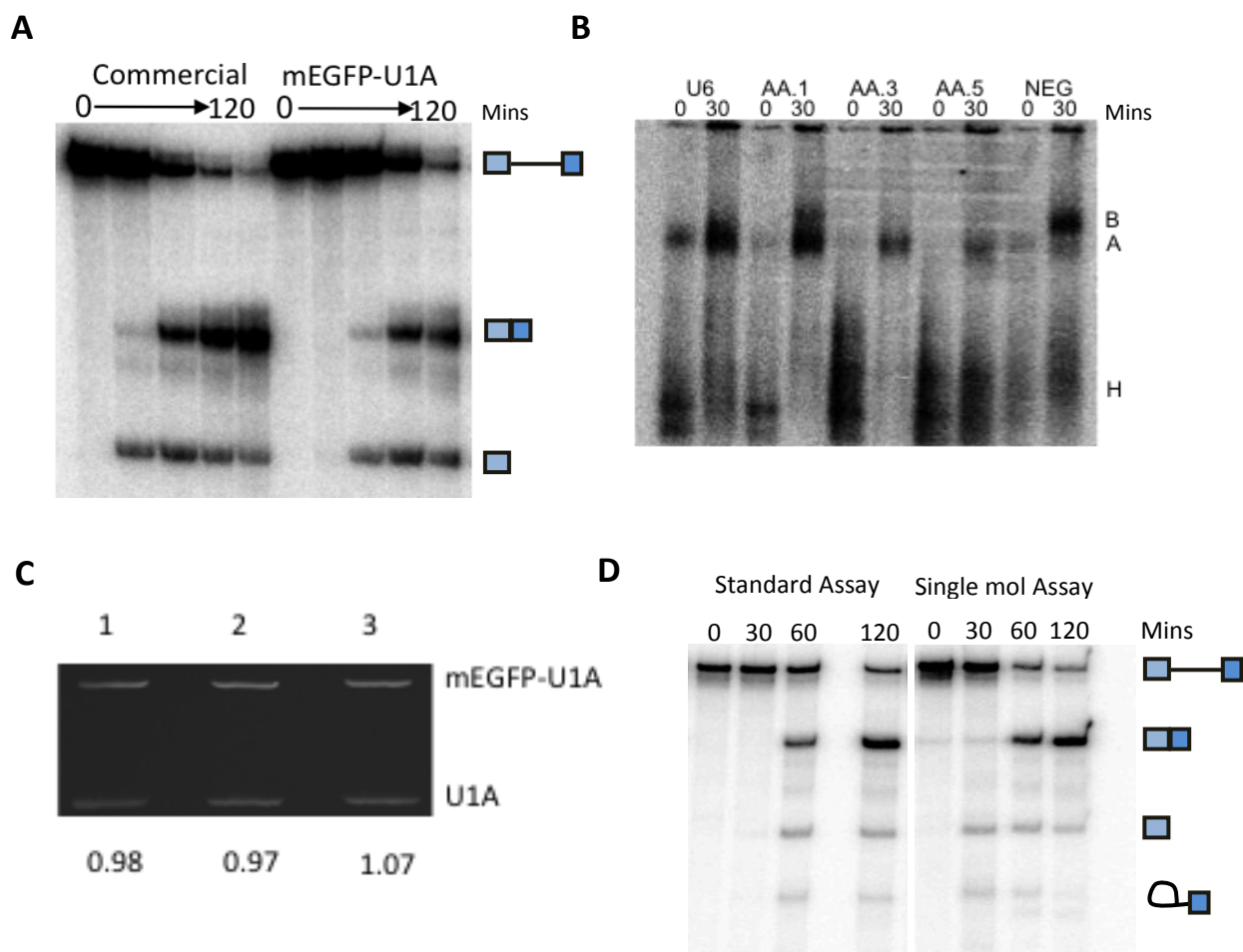


Figure 4.4: Characterization of NE made from mEGFP-U1-A-transfected HeLa cells. (A) Denaturing urea 6 % polyacrylamide gel showing splicing of globin C in commercial HeLa NE or 'home-made' NE containing mEGFP-U1A. Timepoints of 0, 15, 30, 60 and 120 minutes were taken. (B) Stalling of spliceosome formation on 50 nM globin C pre-mRNA in mEGFP-U1-A containing NE, using 1 μ M α -U6 oligonucleotide (Table 2.1) or anacardic acid (AA) at 0.1 mM (AA.1), 0.3 mM (AA.3) or 0.5 mM (AA.5). (C) Fluorescent Western blot of mEGFP-U1-A-containing NE; three samples were run on a denaturing SDS 12 % polyacrylamide gel and the blot was probed using the α U1A(b) antibody followed by the α Mouse fluorescent secondary antibody (Table 2.3). The ratios of mEGFP-U1-A: endogenous U1-A for each lane are shown below the blot. (D) Comparison between efficiency of Globin C splicing under standard splicing conditions and those used for single molecule analysis shown on a denaturing urea 6 % polyacrylamide gel.

experiments, images containing fluorescently labelled pre-mRNA and protein could be moved by the same amount to reveal co-localization.

To determine the signal associated with a single U1 snRNP bound to a piece of RNA, mEGFP-U1-A containing NE was incubated in a splicing reaction. This reaction lacked ATP, and contained a 2' O methyl modified fluorescent oligonucleotide, Anti-U1-Cy5 (Table 4.2), which was complementary to the 5' end of U1 snRNA. This oligonucleotide was modified at the 5' terminus with a Cy5 fluorophore and the 3' terminus with a biotin group. Therefore oligonucleotide-annealed U1 snRNP would be recruited to the surface of the slide just like the pre-mRNAs used in future experiments.

To anneal the oligonucleotide to the U1 snRNA, mEGFP-U1-A containing NE, from which the ATP had been pre-depleted by incubation at 30 °C for 30 minutes, was added to a splicing reaction containing 50 nM Anti-U1-Cy5 oligonucleotide, and then incubated at 30 °C for 15 minutes. Following incubation, the splicing reaction was diluted ready for analysis using the TIRF microscope. Images were obtained from the sample and were then overlaid as described above. These samples showed high levels of co-localization between the red spots, which indicate Cy5, and the green spots, which indicate mEGFP-U1-A (Figure 4.5B). To analyse the number of fluorophores within each co-localized mEGFP-U1-A spot, the position of each co-localized spot on the "sum image" was identified in the "raw image" stack. Once the position was identified, the area occupied by the spot was selected and a z-axis profile was drawn using Image J. This profile shows the intensity of the mEGFP-U1-A spot for each frame

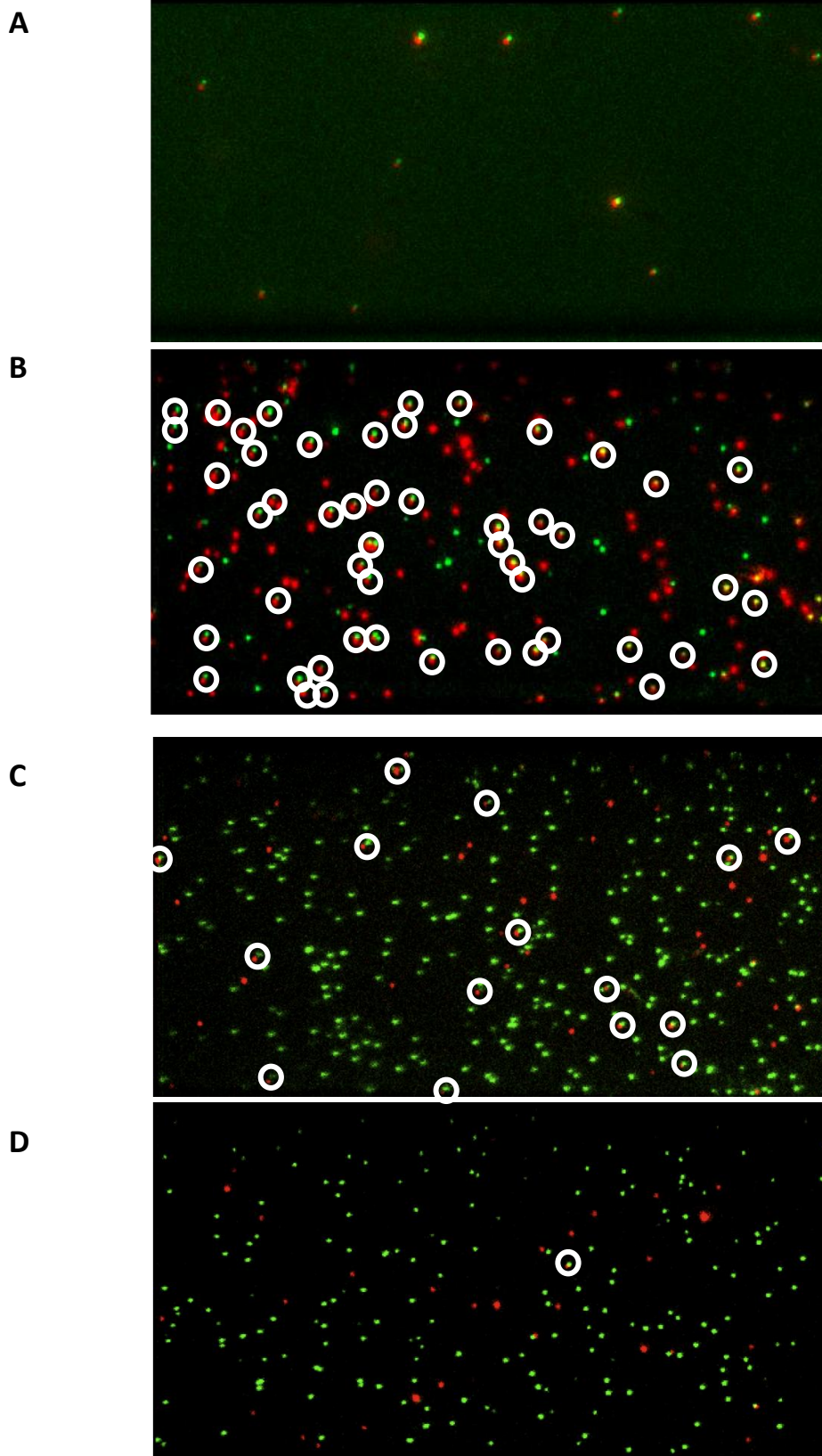


Figure 4.5: Examples of co-localization data. (A) Fluorescent beads being used to determine how far the two images need to be moved in order to assess co-localization. (B) Field showing co-localization between mEGFP-U1-A and the Anti-U1-Cy5 oligonucleotide. White circles indicate co-localized spots; the green channel has been shifted to the right and up a little relative to the red channel to allow observation of both colours. (C) Field showing co-localization between Cy5 oligonucleotide annealed globin C RNA and mEGFP-U1-A. (D) Field showing very low levels of co-localization between Cy5 oligonucleotide-annealed globin M RNA and mEGFP-U1-A.

captured in a graphical format. Since each frame lasts 200ms, this gives a time-resolved profile of the fluorescence intensity for the selected spot. This profile can then be used to determine the number of fluorophores within the spot (for examples of where this technique has been used: Leake et al., 2006; Ulbrich and Isacoff., 2007; Shu et al ., 2007). The number of fluorophores within each spot can be determined because as each fluorophore bleaches, there will be a sudden drop in the overall intensity of the fluorescence detected.

When there is only a single fluorophore in a spot, the signal intensity will drop to the background in one step. Figure 4.6 shows examples of intensity profiles where the intensity profile of the selected spot shows a single step from the highest intensity point to the background. The first 20-30 frames of each intensity profile showed no EGFP fluorescence because, as described in Section 2.3.6, in the first 20-30 frames the sample was excited with the He-Ne 633 nm laser (Table 2.4). To detect the position of the RNA, fluorescence was detected in the mEGFP channel only when the Argon Ion 488 nm laser (Table 2.4) is switched on. The Argon Ion 488 nm laser was kept on until all of the frames had been captured. The background signal is often increased under excitation using the Argon Ion 488 nm laser as seen in Figure 4.6; this is presumably because of a slightly increased background from the fluorophores in solution and an increase in auto-fluorescence from the silica slide.

If there were more than one fluorophore in a spot, then there would be more than one bleaching step, as it is highly unlikely that two fluorophores would bleach at exactly the same time. Intensity profiles with two steps were identified and show the

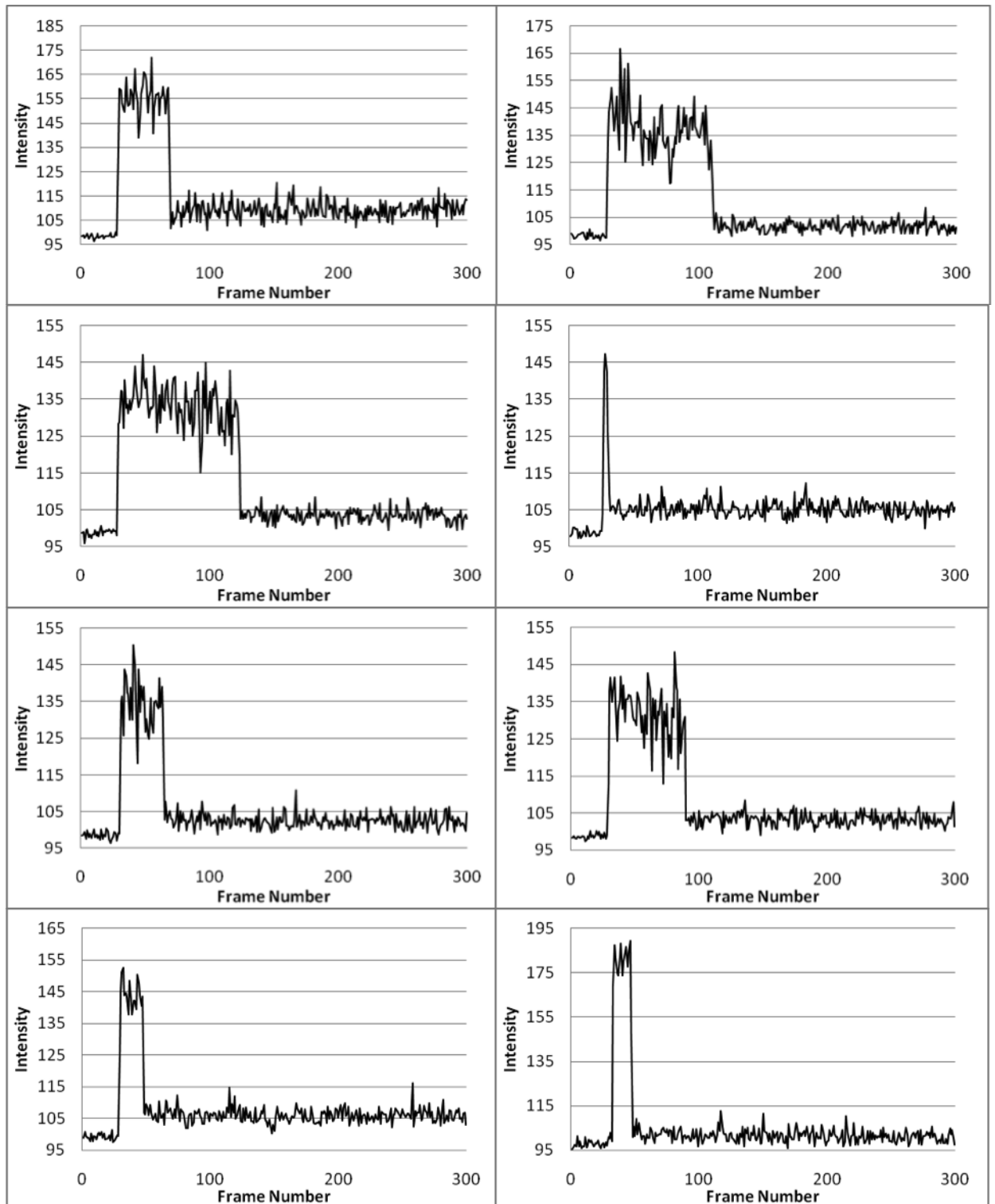


Figure 4.6: Representative single step intensity profiles. Intensity profiles of fluorescence obtained from mEGFP-U1A spots co-localized with Cy5-labelled RNA showing one step, this indicates one fluorophore.

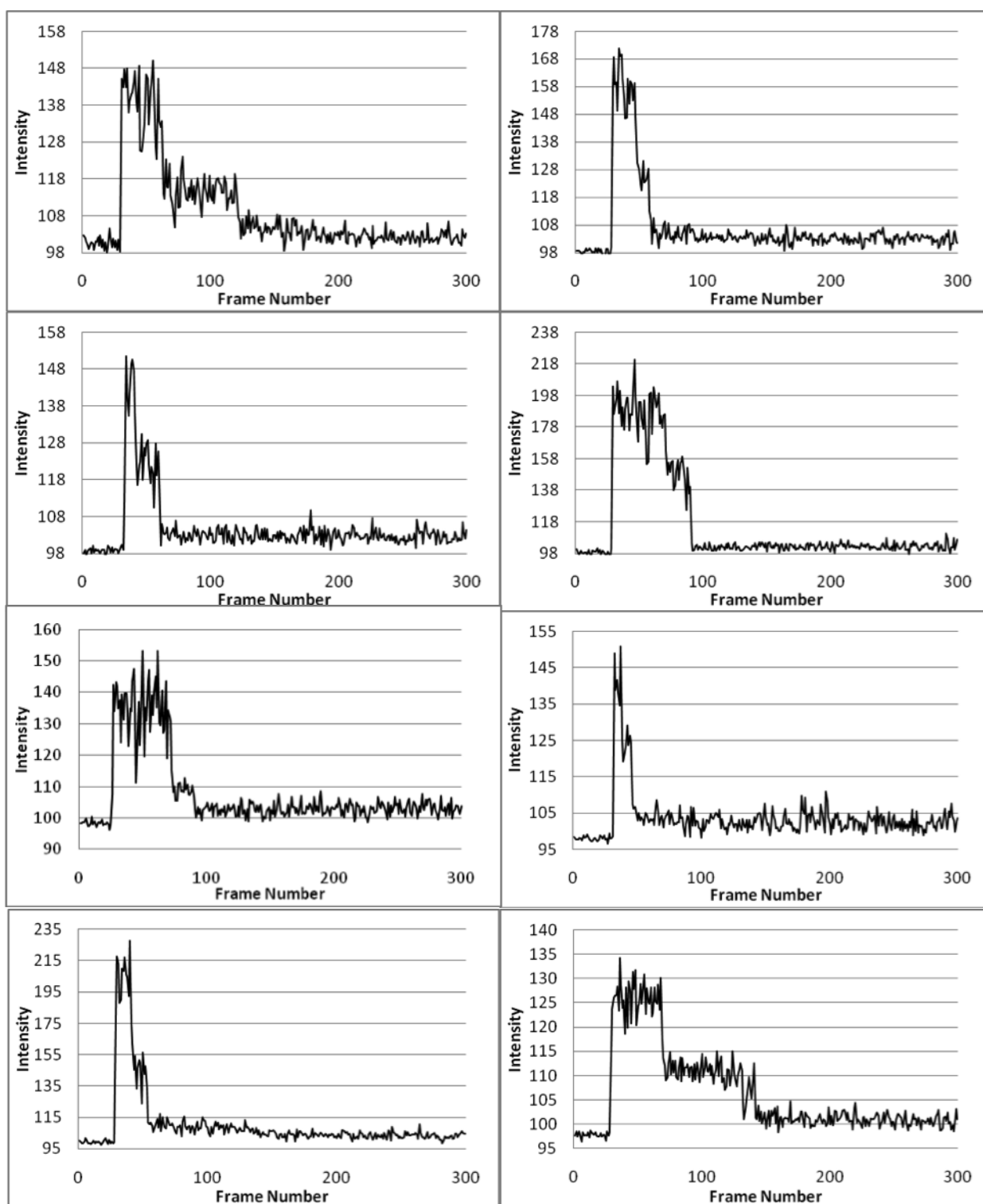


Figure 4.7: Representative two step intensity profiles. Intensity profiles of fluorescence obtained from mEGFP-U1A spots co-localized with Cy5-labelled RNA showing two steps; this indicates two fluorophores.

presence of two fluorophores within the spot; examples of intensity profiles showing this are given in Figure 4.7. Some fluorophores do not show this characteristic bleaching in their intensity profiles and instead produce unidentifiable intensity profiles. Examples of unidentifiable intensity profiles are given in Figure 4.8. These profiles can be generated for several reasons: the fluorophore does not bleach; it bleached but not in a clear step or the intensity of the fluorescence is highly variable. The last group is probably caused by the phenomenon of flickering and blinking fluorophores (Bagshaw and Cherny, 2006). For the purpose of this investigation, any profiles that could not be identified as single step or double step were included within this group. In addition, the small number of profiles that showed 3-4 step bleaching were included in this set. This was because, if any were detected, the number was very low, typically below 5 % of the total number of spots in a data set.

Analysis of the proportions of mEGFP-U1-A spots that co-localized with the Anti-U1-Cy5 oligonucleotide showed that, of the spots that could be assigned as one or two steps, approximately 90 % had a single step intensity profile whilst 10 % had a double step intensity profile (Table 4.3). To determine whether this could indicate that two U1 snRNPs are able to bind this oligonucleotide, expected values for the proportion of one and two step intensity profiles were calculated. The expected values were based on a binomial distribution calculated from the ratio between the endogenous U1-A and transfected mEGFP- U1-A taken from the Western blot shown in Figure 4.4C. Comparison of the expected and observed data sets by Chi-square clearly shows that the observed proportion of intensity profiles with two steps was much lower than would be expected if two U1 snRNPs were present.

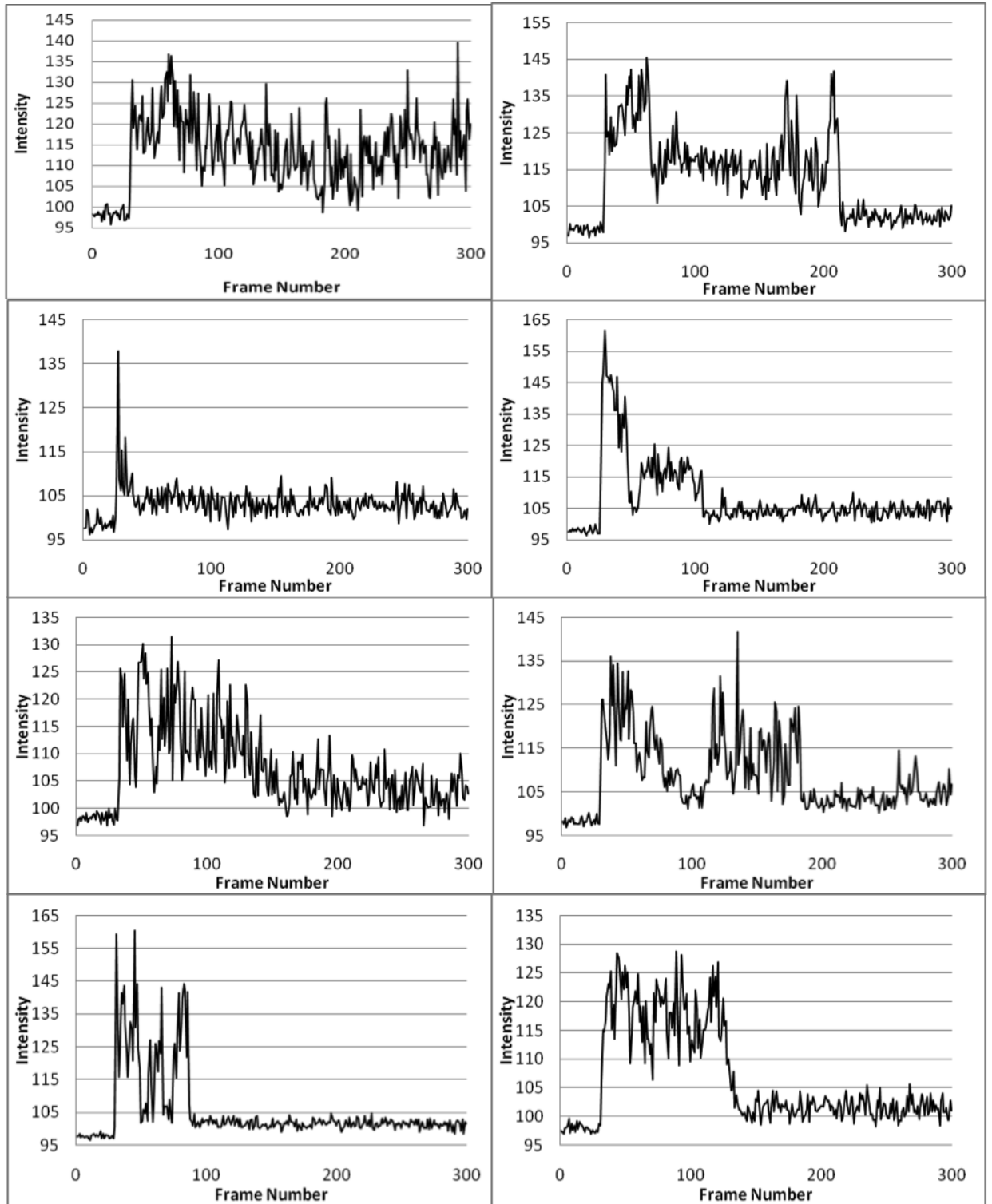


Figure 4.8: Representative unassignable intensity profiles. Intensity profiles of fluorescence obtained from mEGFP-U1A spots co-localized with Cy5-labelled RNA showing no clear step; this indicates an unknown numbers of fluorophores.

Since Anti-U1-Cy5 oligonucleotide was expected to only bind a single U1 snRNP, it was surprising that as many as 10 % of the spots showed two step intensity profiles. An analysis of non-co-localized mEGFP-U1-A spots showed the same proportion of one and two step intensity profiles (Table 4.3). A possible cause is the weak ability of U1-A to homodimerise (Klein Gunnewiek et al., 2000; Varani et al., 2000), leading to a low level of weakly homodimerised U1-A associating with the U1 snRNP.

Sample	Photo-bleaching in one step	Photo-bleaching in two steps	Unassigned	χ^2	P
Anti-U1-Cy5					
Observed	298	39	63		
Expected if 2U1	224	113	-	72.91	<0.001
Non co-localised					
Observed	72	9	24		
Expected if 2U1	54	27	-	18	<0.001
Expected if 1U1	72	9	-	0	>0.95

Table 4.3: Quantification of the number of mEGFP-U1-A in spots either co-localized with the Anti-U1-Cy5 oligonucleotide or non co-localized

Analysis of the mEGFP-U1-A spots that were co-localized with fluorescent oligonucleotide-labelled globin C showed a similar proportion of single and double step intensity profiles (Table 4.5). Comparison between the expected values for two U1 snRNPs and the observed values, again showed there could not be two U1 snRNPs present. Since the proportions were very similar to the intrinsic values for the U1 snRNPs in Table 4.3 it follows that globin C was bound by only one U1snRNP. 26 % of the globin C molecules were associated with mEGFP-U1-A, whereas only 3 % of the globin M molecules showed such co-localization (Table 4.4; Figure 4.5 C and D). Thus, the majority of co-localization depends on a bona fida 5' SS.

Sample	Total Cy5 spots	Co-localized Cy5 spots	Percentage co-localization
globin C –ATP	346	91	26%
globin C +U6	188	27	14%
globin C +AA	264	41	16%
globin M –ATP	112	3	3%
globin M +U6	115	3	3%
C24C –ATP	127	41	32%
C24C +U6	232	12	5%
C24C +AA	259	13	5%
C174C –ATP	195	66	34%
C174C +U6	182	47	26%
C174C +AA	246	91	37%
C174M –ATP	242	59	24%
M174C –ATP	396	91	23%
Ad1CC –ATP	216	76	35%
Ad1CC +U6	116	54	46%
Ad1CC +AA	363	97	27%
CEC –ATP	233	106	45%
CEC +U6	272	55	20%
CEC +AA	202	64	32%
C174C Δ 3' –ATP	210	90	43%
Ad1CC +U6 –ATP	190	86	45%
Ad1CC +U6 +ATP	236	73	30%

Table 4.4: Approximate efficiency of co-localization between mEGFP-U1-A and Cy5 labelled pre-mRNA.

4.5 Two U1 snRNPs probably associate with pre-mRNAs containing two 5' SS under E complex conditions

TIRF microscopy experiments with the C24C and C174C pre-mRNAs, which contain two 5' SS, showed markedly different to the proportions of co-localized spots with single or double step intensity profiles. C24C, which splices inefficiently to both consensus 5' SS (Figure 4.3A), had a higher proportion of two step intensity profiles; when the proportions were compared with those expected for occupation of the pre-mRNA by one or two U1 snRNPs, it was found that they were concomitant with the expected

Sample	Photo-bleaching in one step	Photo-bleaching in two steps	Unassigned	χ^2	P
--------	--------------------------------	---------------------------------	------------	----------	---

values for the binding of two U1 snRNPs (Table 4.5). C174C, which splices efficiently but only to the downstream 5' SS, had similar proportions of one and two step intensity profiles (Table 4.5).

As controls, the C174M and M174C pre-mRNAs were also tested. These pre-mRNAs, which contain the same 174nt insertion as the C174C but lacked one of the 5' SS, had proportions of one and two step intensity profiles similar to those seen with globin C rather than C174C (Table 4.5). This showed that the increase in two step intensity profiles in C174C was due to the presence of a second 5' SS and not due to the inserted sequence. The Ad1CC RNA data also showed agreement with the results of the globin series of pre-mRNAs and were concomitant with the presence of two U1 snRNPs on each molecule (Table 4.5).

Overall, this data suggests that two U1 snRNPs can associate with the pre-mRNAs containing two strong 5' SS. On this basis the hypothesis that affinity is the sole determinant of 5' SS selection (Figure 1.3) can be rejected. The results agree with the model wherein the two 5' SS are recognised independently, with both affinity and position having a role in 5' SS selection (Figure 1.4).

A key hypothesis of this model is that if the concentration of free U1 snRNP within the NE was reduced, the relative use of the two 5' SS would be altered. This is because a reduction in the amount of free U1 snRNP should lead to a reduction of the proportion of pre-mRNAs that are bound by two U1 snRNPs and would lead to an increase in the proportion of pre-mRNAs bound by one U1 snRNP at either the upstream or downstream 5' SS. On these transcripts, the 5' SS that is occupied would be used

globin C					
Observed	93	12	22		
Expected if 2 U1	70	35	-	23.15	<0.001
Expected if 1 U1	93	12	-	0	>0.95
C24C					
Observed	71	38	28		
Expected if 2 U1	72	37	-	0.04	>0.95
Expected if 1 U1	96	13	-	54.59	<0.001
C174C					
Observed	75	35	24		
Expected if 2 U1	73	37	-	0.16	>0.9
Expected if 1 U1	97	13	-	42.22	<0.001
C174M					
Observed	96	14	21		
Expected if 2 U1	73	37	-	21.54	<0.001
Expected if 1 U1	97	13	-	0.14	>0.95
M174C					
Observed	95	16	30		
Expected if 2 U1	74	37	-	18.33	<0.001
Expected if 1 U1	98	13	-	0.78	>0.6
Ad1CC					
Observed	69	35	22		
Expected if 2 U1	69	35	-	0	>0.95
Expected if 1 U1	92	12	-	49.83	<0.001

Table 4.5: Proportion of mEGFP-U1-A spots which bleached in one and two steps when under Complex E conditions.

(Figure 1.4) and therefore an increase in the relative level of upstream 5' SS use would be expected. To test this hypothesis, an experiment was carried out in which the α -U1 oligonucleotide (Table 2.1) was titrated into a splicing reaction containing the C175G pre-mRNA. The C175G pre-mRNA was used instead of the C174C pre-mRNA because it still showed a strong position effect (Figure 4.9, sample O; Cunningham et al., 1991), but had a slightly weaker downstream 5' SS sequence. Therefore the α -U1 oligonucleotide was expected to be able to compete with the 5' SS for the binding of U1 snRNP more efficiently. To allow the binding of U1 snRNP by the α -U1 oligonucleotide, NE was pre-incubated with the given concentration of oligonucleotide for 15 minutes prior to addition of the NE to a standard splicing reaction. Splicing reactions for each concentration of α -U1 oligonucleotide were carried out in triplicate.

The results of this experiment are displayed in Figure 4.9 and clearly show that with increasing concentrations of the α -U1 oligonucleotide, the proportion of mRNA formed using the upstream 5' SS increases; this is especially noticeable when the samples incubated with 0.2 μ M were compared to the 0 μ M sample. It is also noticeable that the overall level of splicing is reduced, probably due to there being less free U1 snRNP in the NE.

4.6 Loss of one U1 snRNP from pre-mRNAs containing two 5' SS by A complex

Experiments have shown 5' SS selection to be committed by A complex (Kotlajich et al., 2009). To see whether this is reflected in any changes in the number of U1 snRNPs associated with pre-mRNAs containing two 5' SS, mEGFP-U1-A containing NE was pre-incubated with either anacardic acid or the α -U6 oligonucleotide before the NE was added to a splicing reaction. The reaction contained 50 nM fluorescent oligonucleotide-labelled pre-mRNA and was incubated at 30 $^{\circ}$ C for 30 minutes to allow complex formation.

Under these conditions, globin C had very similar proportions of one and two step intensity profiles when compared to the pattern seen under E complex conditions. This showed that the number of U1 snRNPs present on the globin C pre-mRNA has not altered between the E and A complexes (Table 4.6). Analysis of globin M under A complex conditions again showed very low levels of co-localization showing the co-localization observed to be specific (Table 4.4).

In contrast, when C174C was analysed under A complex conditions, the proportion of intensity profiles which showed two steps was greatly reduced in comparison to those

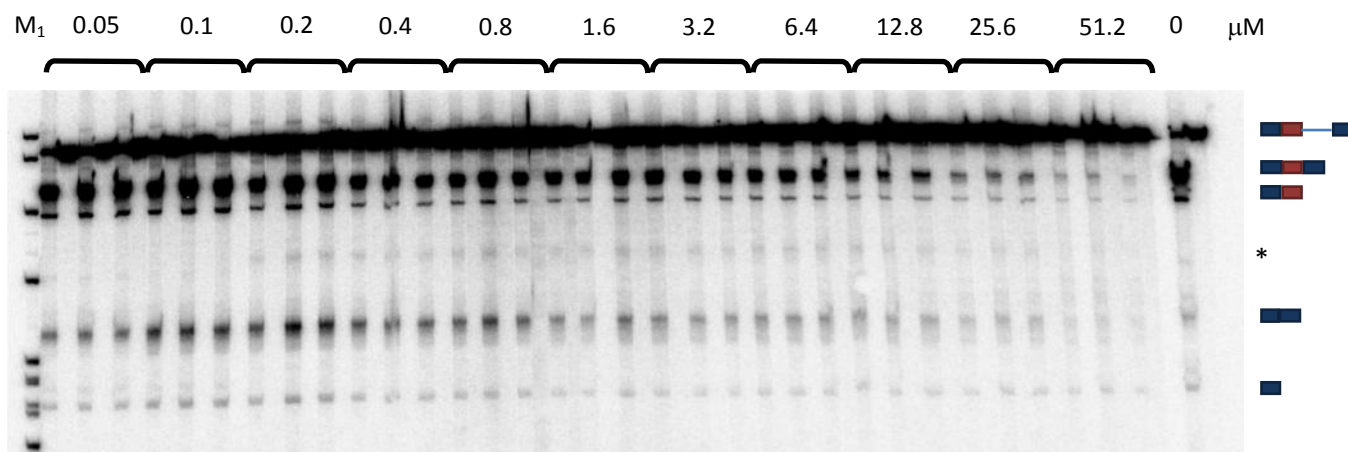


Figure 4.9: Altering the 5' SS selection on the C175G pre-mRNA by addition of the α -U1 oligonucleotide (Table 2.1). Each concentration of oligonucleotide was tested in triplicate with time points of 0 and 120 minutes and is shown on a denaturing urea 6 % polyacrylamide gel. Concentration '0' has had no oligonucleotide added. The RNA product labelled (*) indicates probable cryptic 5' SS use.

with C174C under E complex conditions. Comparison between the observed values and expected values for occupancy of the pre-mRNA by one or two U1 snRNPs indicated that under A complex conditions, C174C only has one U1 snRNP attached (Table 4.6). This highly interesting result suggested that during the conversion from E complex to A complex the U1 snRNP associated with the unused 5' SS is removed from the pre-mRNA. The U1 snRNP that remains must be bound to the 5' SS that is used and not to the unused 5' SS, as globin C is still bound by U1 snRNP under A complex conditions.

Experiments carried out on the Ad1CC substrate under A complex conditions allowed the same conclusion to be drawn, as a drop in the proportion of two step intensity profiles was observed (Table 4.6). Again, comparison of the observed values with the expected values for occupancy by one or two U1 snRNPs showed that occupancy by one U1 snRNP was correct.

One possible reason for this loss of one U1 snRNP was that the RNA was being cut. To assess whether this was occurring, Ad1CC pre-mRNA was annealed to two oligonucleotides. The first of these was oligo 9 which was usually used with Ad1CC, the second was the Ad1-intron oligonucleotide (Table 4.2), which is modified at the 5' terminus with a Cy-3 fluorophore and anneals to the Ad1CC pre-mRNA within the intron. Annealing of both these oligonucleotides to the Ad1CC pre-mRNA allowed assessment of the level of degradation that occurred under different conditions. Analysis of the percentage of Cy5 spots that were co-localised with Cy3 spots under E complex (-ATP) or A complex conditions (+AA/+U6) showed that there was very little difference between the conditions (Table 4.7). This indicated that the drop in the

Sample	Photo-bleaching in one step	Photo-bleaching in two steps	Unassigned	χ^2	P
globin C +AA					
Observed	69	7	16		
Expected if 2 U1	50	26	-	21.10	<0.001
Expected if 1 U1	67	9	-	0.5	>0.75
globin C +U6					
Observed	66	7	17		
Expected if 2 U1	48	25	-	19.71	<0.001
Expected if 1 U1	65	8	-	0.14	>0.9
C174C +AA					
Observed	86	12	28		
Expected if 2 U1	65	33	-	20.15	<0.001
Expected if 1 U1	87	11	-	0.10	>0.95
C174C +U6					
Observed	79	10	7		
Expected if 2 U1	59	30	-	20.11	<0.001
Expected if 1 U1	79	10	-	0	>0.95
Ad1CC +AA					
Observed	95	14	41		
Expected if 2 U1	72	37	-	21.64	<0.001
Expected if 1 U1	96	13	-	0.09	>0.9
Ad1CC +U6					
Observed	50	6	14		
Expected if 2 U1	37	19	-	13.46	<0.001
Expected if 1 U1	50	6	-	0	>0.95
C174C Δ3' + U6					
Observed					
Expected if 2 U1	70	26	44		
Expected if 1 U1	64	32	-	1.69	>0.4
	85	11	-	23.10	<0.001

Table 4.6: Proportion of mEGFP-U1-A spots which bleached in one and two steps when under Complex A conditions.

Sample	Cy5 spots analysed	% colocalised with Cy3
Ad1CC Dual -ATP	368	80
Ad1CC Dual + AA	510	77
Ad1CC Dual + U6	146	78

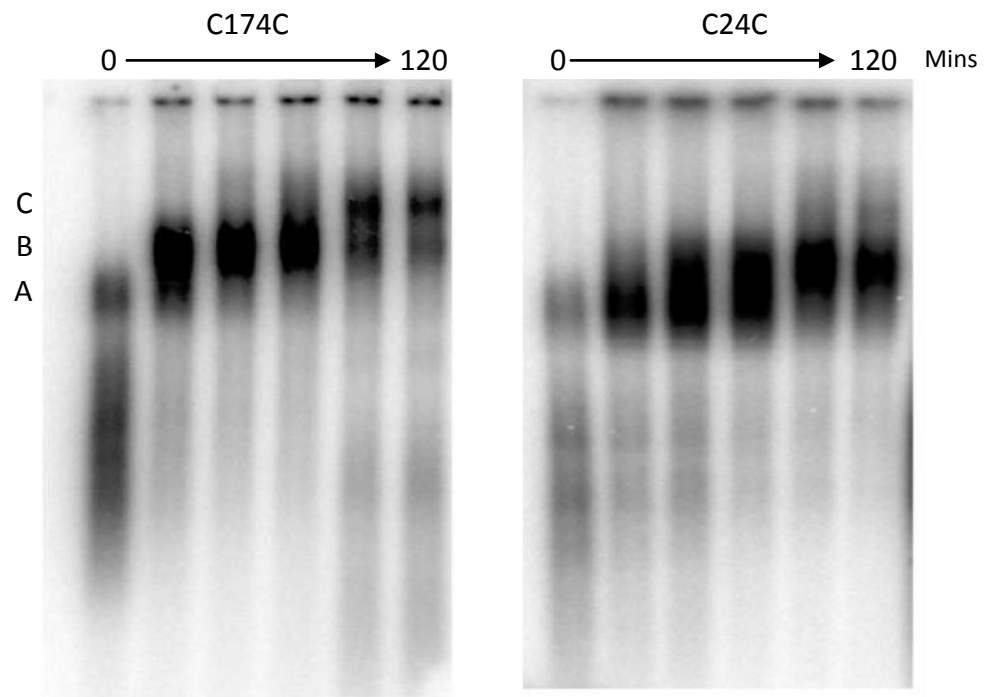
Table 4.7: Co-localization between Cy5- and Cy3- labelled oligonucleotides attached to Ad1CC pre-mRNA.

number of U1 snRNPs bound to the pre-mRNA was not due to endonuclease digestion of the pre-mRNA. Degradation by exonucleases cannot be ruled out by this experiment as the oligonucleotides are likely to block progress due to their high stability.

In contrast to the C174C and Ad1CC pre-mRNAs, the C24C pre-mRNA co-localized with mEGFP-U1-A very inefficiently under A complex conditions (Table 4.4). Therefore the number of U1 snRNPs bound to C24C under A complex conditions could not be determined.

To investigate what could have caused this, the C24C and C174C pre-mRNAs were analysed by native LMP agarose gel electrophoresis, since it was possible that complex formation on C24C was inhibited. For this experiment, the same amounts of radioactively labelled C24C and C174C, as assessed by Cherenkov counting, were added to their respective splicing reactions to allow comparison. The time-points in which C174C the A, B and C complexes accumulated (Figure 4.10A) were very similar to those of globin C (Figure 4.10B) suggesting that, as expected, on C174C spliceosome accumulation was not inhibited. In contrast, although C24C showed similar accumulation of A complex, the transition to B complex appeared to be severely inhibited when compared to C174C (Figure 4.10A). C174C was almost entirely in B complex by the 15 minute time-point whilst C24C did not reach this point until 60 minutes. The complexes formed on C24C appeared as a smear. To confirm that the A complex to B complex transition was being observed, C24C was incubated in NE that had been pre-incubated with 1 μ M α -U6 oligonucleotide (Table 2.1). Analysis of the complexes showed that only the lower complex was formed, confirming that this is A

A



B

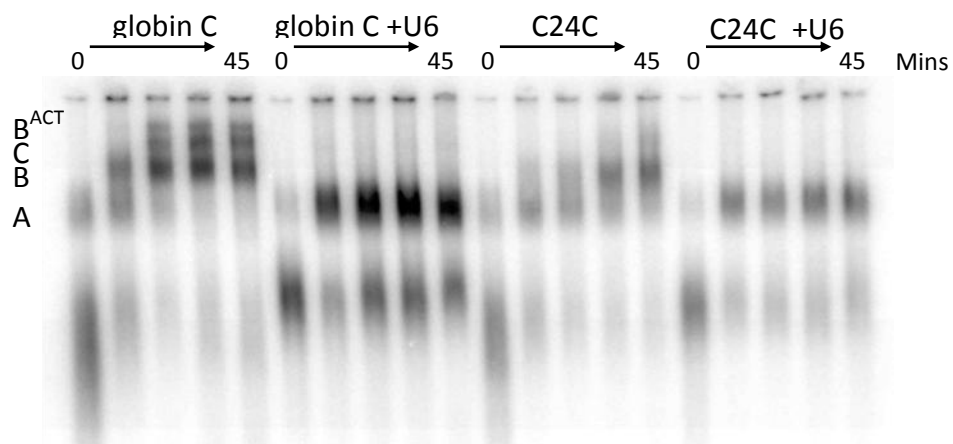


Figure 4.10: Spliceosome assembly on the C24C pre-mRNA. (A) Native 2 % LMP agarose gel of splicosomal complexes formed on the C174C and C24C RNAs. Time points of 0, 5, 15, 30, 60 and 120 minutes were taken. **(B)** Native 2 % LMP agarose gel of splicosomal complexes formed on globin C and C24C RNA in the presence (+U6) or absence of the α -U6 oligonucleotide, with time points of 0, 5, 15, 30 and 45 minutes taken.

complex. This data indicates that the inhibition of C24C splicing may occur between Complexes A and B. One hypothesis that may explain this data is that on C24C both U1 snRNPs are removed by A complex due to their close proximity, leading to inhibition of splice site selection. The inhibition of the transfer from A complex to B may be because, without the U1 snRNP, the spliceosome is unable to define the 5' SS and therefore the low level of splicing observed is because of other mechanisms of 5' SS selection. This is possibly through recognition of the 5' SS by the U5 or U6 snRNPs, which are known to be able to influence 5' SS selection (Cortes et al., 1993; Hwang and Cohen, 1996).

The experiments so far described have been carried out using complexes stalled at a single point in spliceosome assembly. As an extension of this observation, the ability of a stalled E complex to release one of the U1 snRNPs when ATP was added was investigated. To do this ATP was depleted from the NE in the presence of the α -U6 oligonucleotide and this NE was then added to a splicing reaction lacking ATP and Phosphocreatine. It was incubated at 30 °C for 30 minutes to allow E complex assembly and a sample was then taken for analysis by TIRF microscopy. The remainder of the splicing reaction was then incubated, in the presence of 1.36 mM ATP and 19 mM Phosphocreatine, for a further 30 minutes at 30 °C to allow A complex formation; after this incubation a second sample was taken for analysis by TIRF microscopy. Analysis of the data obtained from this experiment shows that, prior to the addition of ATP and Phosphocreatine, the proportions of spots that bleached in one or two steps was in agreement with the expected values for occupation by two U1 snRNPs (Table 4.8). In contrast, when the sample to which ATP and Phosphocreatine

Sample	Photo-bleaching in one step	Photo-bleaching in two steps	Unassigned	χ^2	P
Ad1CC +U6 pre-ATP addition					
Observed	60	30	54		
Expected if 2 U1	60	30	-	0	>0.95
Expected if 1 U1	80	10	-	45	<0.001
Ad1CC +U6 post-ATP					
Observed	72	15	50		
Expected if 2 U1	58	29	-	12.01	<0.005
Expected if 1 U1	77	10	-	2.82	>0.2

Table 4.8: Proportion of mEGFP-U1-A spots which bleached in one and two steps when co-localized with Ad1CC pre and post ATP addition.

had been added was analysed, the proportions of one and two step bleaching were no longer in agreement with the expected values for occupation by two U1 snRNPs and instead were closer to the values expected for a single U1 snRNP. This suggests that the loss of the U1 snRNP associated with the unused 5' SS can be triggered by the addition of ATP. However, it cannot be ruled out that the complexes seen in the sample after ATP and phosphocreatine addition were new complexes, i.e. that the complexes assembled under E complex conditions have been disassembled and new ones assembled upon addition of ATP. However, this seems unlikely, considering the evidence that purified E complex can be converted into the ATP-dependent spliceosome (Michaud and Reed, 1991).

4.7 Loss of U1 snRNP is dependent on the presence of a 3' SS downstream and is inhibited by the presence of a 3' SS between the two 5' splice sites

To assess the effect of the 3' SS region of C174C on the number of U1 snRNPs associated with the pre-mRNA, an RNA was made lacking the 3' end. This RNA, C174CΔ3', was analysed by TIRF microscopy under A complex conditions to see whether a 3' SS was needed for the loss of the U1 snRNP associated with the unused 5'

SS. The analysis showed that C174CΔ3' probably had two U1 snRNPs associated under A complex conditions (Table 4.6); this showed that the 3' end of C174C, probably the 3' SS, was needed for release of the unused U1 snRNP from the 5' SS and thus intron definition was required.

One question that remained was how the spliceosome ensured, that for a pre-mRNA containing two introns, both introns were spliced. Why would the U1 snRNP occupying the downstream 5' SS not cause removal of the U1 snRNP associated with the upstream 5' SS? It was hypothesised that the 3' SS region separating the two 5'splice sites must inhibit the loss of the U1 snRNP associated with the upstream 5' SS. To test this, a construct was made in which the globin C exon 2-intron 2-exon 3 unit was duplicated, named CEC.

To make CEC, a shortened fragment of globin C was amplified by PCR using the oligonucleotides "BGE3/E2 repeat" and "BGE3 3' rep" (Table 4.1). This fragment was truncated at the 5' end so that it did not contain the florescent-oligonucleotide annealing region for Bg-Cy5-5'. The BGE3/E2 repeat oligonucleotide was designed to have a 20nt region of overlap with the 3' end of exon 3; this allowed the 5'end of the shortened fragment to anneal to the 3' end of the full length construct. To create CEC from these two fragments, a PCR reaction was set up containing both fragments and the oligonucleotides, P16 and Ad1-trx R (Table 2.1). P16 annealed to the same region of the globin C construct as the Bg-Cy5-5' fluorescent oligonucleotide and thus was only able to anneal to the full-length globin C and not the shortened fragment. Ad1-trx-R was used for the reverse strand as the normal reverse strand oligonucleotides (P17, βG E3 no 5' SS) could not be used as both fragments contained complementary

sequences and only the single intron construct would have been produced. Ad1-trx-R could be used as the BGE3 3' rep oligonucleotide used to amplify the truncated fragment of globin C contained a complementary sequence. Amplification using the P16 and Ad1-trx R oligonucleotides could only produce a construct where the globin C unit was duplicated. Once cloned and sequenced the CEC construct was PCR amplified using the P16 and β G E3 no 5' SS oligonucleotides, ready for transcription. The complementary sequence for the β G E3 no 5' SS oligonucleotide within the central exon had been removed as part of its creation. Transcription of this construct was not clean and produced multiple bands; to analyse this pre-mRNA on the single molecule level, pre-mRNA produced by cold transcription had to be purified by ultra violet shadowing (Section 2.3.2) instead of the standard procedure in Section 2.3.1.

Analysis of the products formed by CEC when spliced in a standard splicing reaction showed that there was no exon skipping, with the major product being an mRNA consisting of all three exons (Figure 4.11A). Analysis of the intermediates formed indicated that there was not a preferential order of intron removal, with all possible 5' exon intermediates being detectable. Bands were assigned provisionally based upon migration relative to markers. To confirm the assignments, shift assays were done using the Bg-Cy5-5' fluorescent oligonucleotide, which indicated which of the bands contained the 5' terminus of the transcript. To anneal the small amount of radiolabelled CEC pre-mRNA to the BG-Cy5-5' oligonucleotide, an annealing reaction was set up containing an entire hot transcription and 50 nM BG-Cy-5' in Annealing buffer (Section 2.5.1) and was heated to 80 °C for 2 minutes before incubation at 4 °C for 10 minutes and addition to the splicing reaction. The results are shown in Figure

4.11B. Identification of bands as lariats was through the observation of altered migration when run on different percentage polyacrylamide denaturing urea gels (data not shown).

Analysis of the complexes formed on the CEC pre-mRNA by native 2 % LMP agarose gel electrophoresis showed that a series of very large complexes were formed (Figure 4.11C). The lack of smaller complexes suggests that the spliceosome is able to form over both introns at the same time, as previously reported (Christofori et al., 1987). To confirm that the complexes detected were spliceosomes forming over the intron, the central exon was cut using the α 3EDNA oligonucleotide (Table 2.1) and RNaseH, as described in Section 2.1.6. This oligonucleotide was designed to anneal over the artificial join created in the central exon so that it would cut uniquely. Analysis of the complexes formed after treatment with RNaseH and the α 3EDNA oligonucleotide, showed that the CEC pre-mRNA forms complexes, which are the same size as globin C (Figure 4.12A). Interestingly the rate of complex formation did not appear to be enhanced or inhibited by the presence of a second intron. To confirm efficient and specific digestion of the CEC pre-mRNA in the central exon, the RNA products of the RNaseH digestion were purified and run on a denaturing urea 5 % polyacrylamide gel (Figure 4.12B).

To confirm that the CEC pre-mRNA formed a discrete complex when stalled at A complex, the CEC pre-mRNA was added to splicing reactions that had been inhibited using either anacardic acid or the α -U6 oligonucleotide (Figure 4.12C). Treatment with the α U6 oligonucleotide led to stalling at a single complex on CEC that was significantly larger than the complex formed on globin C, suggesting that it was a double A

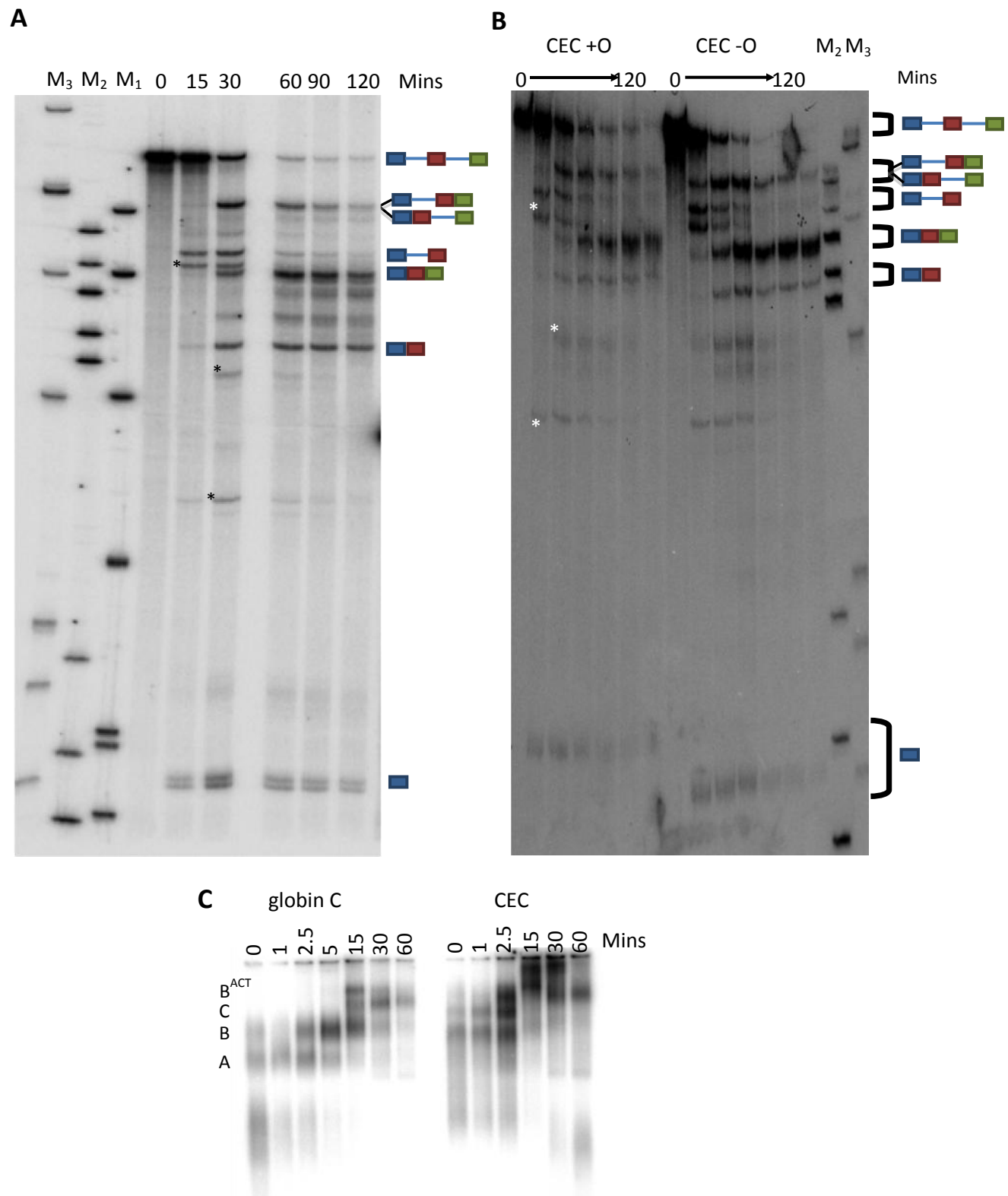


Figure 4.11: Spliceosome assembly on CEC pre-mRNA. (A) Denaturing urea 5 % polyacrylamide gel with 30 % formamide showing the products produced by splicing of CEC, '*' indicates lariats. (B) Denaturing urea 6 % polyacrylamide gel with 30 % formamide showing splicing of CEC either annealed to the BG-Cy5-5' oligonucleotide (+O) or not annealed (-O), white asterix's indicates lariats, time-points were 0, 15, 30, 45, 60, 90 and 120 minutes. (C) Native 2 % LMP agarose gel of splicosomal complexes formed on globin C and CEC.

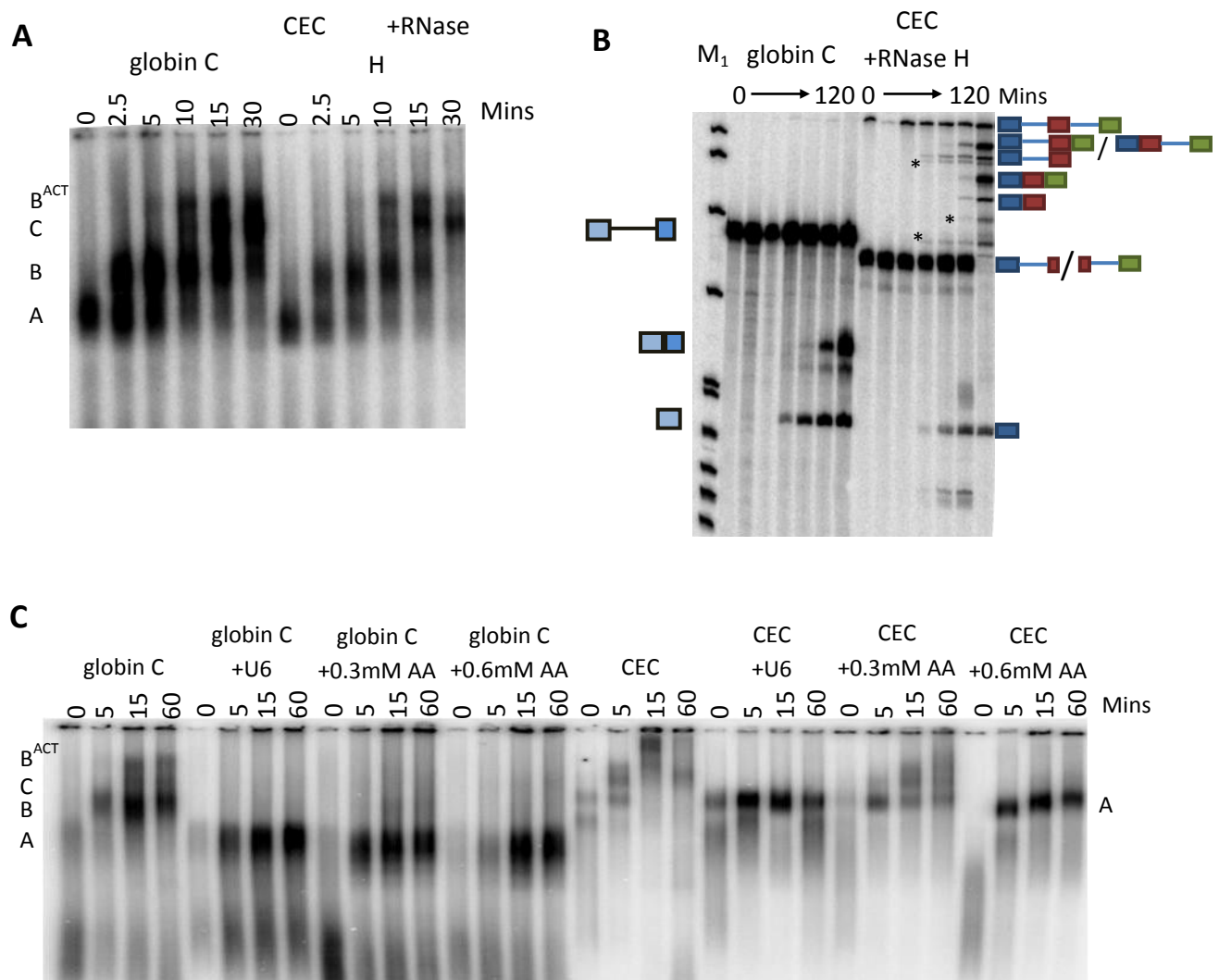


Figure 4.12: Spliceosome assembly on CEC pre-mRNA. (A) Native 2 % LMP agarose gel of spliceosomal complexes formed on globin C and CEC which had been digested with the α 3EDNA oligonucleotide (+RNase H; Table 2.1). **(B)** Denaturing urea 5 % polyacrylamide with 30 % formamide showing the RNA products produced by the samples run in Figure 4.12A, '*' indicates lariats, band produced by cleavage of central exon indicated by half constructs. **(C)** Native 2 % LMP agarose gel showing stalling of spliceosomal assembly on globin C and CEC; U6 refers to the addition of 1 μ M α -U6 oligonucleotide, AA refers to the addition of anacardic acid.

complex. In contrast, splicing of CEC using extract that had been pre-incubated with 0.3 mM anacardic acid led to only an attenuation of complex progression and not stalling at A complex; spliceosome formation on globin C was clearly stalled at A complex under these conditions. To stall spliceosome assembly at A complex on the CEC pre-mRNA, the concentration of anacardic acid had to be increased to 0.6 mM (Figure 4.12C, CEC +0.6 mM AA).

When TIRF experiments were carried out using the CEC pre-mRNA in the absence of ATP, the observed proportions of one and two step intensity profiles are consistent with the binding of two U1 snRNPs to the pre-mRNA. Analysis of samples under A complex conditions also showed the same proportions (Table 4.9). This shows that when pre-mRNAs with two separate introns are incubated under A complex conditions, the U1 snRNP associated with the upstream 5' SS is not lost, presumably due to the presence of a 3' SS between the two 5'splice sites.

Sample	Photo-bleaching in one step	Photo-bleaching in two steps	Unassigned	χ^2	P
CEC -ATP					
Observed	65	36	32		
Expected if 2 U1	67	34	-	0.17	>0.9
Expected if 1 U1	89	12	-	54.47	<0.001
CEC +U6					
Observed	69	31	34		
Expected if 2 U1	66	34	-	0.4	>0.75
Expected if 1 U1	88	12	-	34.19	<0.001
CEC +AA					
Observed	67	29	33		
Expected if 2 U1	64	32	-	0.42	>0.75
Expected if 1 U1	85	11	-	33.27	<0.001

Table 4.9: Proportion of mEGFP-U1-A spots which bleached in one and two steps when co-localized with CEC.

Summary

The work described in this chapter was designed to answer the simple but intriguing question: how many U1 snRNPs can bind a pre-mRNA containing multiple 5'splice sites? From this work, we have determined that a pre-mRNA with two 5'splice sites can be bound by two U1 snRNPs. Interestingly, between the E and A complexes, the U1 snRNP bound to the unused 5' SS is removed by an unknown mechanism. This appears to occur only when the two 5'splice sites are sufficiently separated; if they are not separated fully then both U1 snRNPs are lost and splicing is inhibited. The loss of the U1 snRNP bound to the unused 5' SS also requires the presence of a 3' SS downstream, whilst a 3' SS between two 5'splice sites leads to an inhibition of U1 snRNP removal.

Chapter 5. U2AF65 and its interactions with pre-mRNA

5.1 Preliminary detection of co-localization between mEGFP-U2AF65 and a globin pre-mRNA

5.2 Functional characterization of U2AF65 and U2AF35 fluorescent fusion proteins

5.3 Co-localization between mEGFP-U2AF65 and a 'negative control' pre-mRNA

5.4 Analysis of the 24-36 U/C pY tract of globin C

5.5 An extensive analysis of co-localization efficiencies between mEGFP-U2AF65 and globin C derivatives

5.6 Functional characterization of other fluorescent fusion proteins

Introduction

The spliceosomal factor, U2AF, is required for the recognition of the 3' SS of an intron. The larger subunit of U2AF, U2AF65 recognises the pY tract situated upstream of the 3' SS AG dinucleotide (Singh et al., 1995), whilst U2AF35 recognises the AG dinucleotide (Wu et al., 1999; Zorio and Blumenthal, 1999; Merendino et al., 1999). As discussed in Section 1.5, the pY tract of an intron can vary both in sequence composition and length. The mechanism by which U2AF65 is able to bind and recognise the highly varied pY tracts found in the genome has been the subject of many studies. It has been suggested that recognition of variable pY tracts by U2AF65 is because U2AF65 has multiple modes of binding. This hypothesis appeared to explain how pY tracts with poor sequence composition could be recognized: it was proposed the U2AF65 sampled interactions until it found a suitable binding site. It also gave an explanation for the effect of pY tract length: by increasing the length of a pY tract the number of possible sites is increased and so the probability of binding is also increased (Banerjee et al., 2003). However, as discussed in Section 1.5, an alternative interpretation of experiments published later (Banerjee et al., 2004), suggested that there may be cases where more than one U2AF65 can bind a pY tract.

The number of U2AF65 and U2AF35 that could be bound to a long pY tract was investigated using the pY tract from the rat α -tropomyosin exon three. In the rat α -tropomyosin gene, the second and third exons are mutually exclusive (Wieczorek et al., 1988; Smith and Nadal-Ginard, 1989); this mutual exclusivity is enforced by the close proximity of the 5' SS of exon 2 and the BPS of exon 3. The relative positions of the two exons mean that exon two and exon three cannot be spliced to each other:

either exon two or three is spliced to exons one and four (Smith and Nadal-Ginard, 1989; Mullen et al., 1991). Exon three has strong splicing signals and therefore is the preferred exon in most cell types (Mullen et al., 1991). This choice of exon is altered in smooth muscle cells, where exon two is used, this effect is primarily due to repression of exon three (Gooding et al., 1994; Perez et al., 1997). The strong splicing signals possessed by exon three include a strong pY tract of 50 nt (Smith and Nadal-Ginard, 1989); the length of the pY tract has been found to play an important role in the choice of exon. Only 17 nts of the pY tract are required for splicing in the absence of exon two, but 39 nts are required for use of exon three when it is in competition with exon two (Smith et al., 1989; Mullen et al., 1991). Given the length of the pY tract of this exon, it was the ideal substrate to assess the number of U2AF65 that can be bound to a pre-mRNA.

The regions around this exon had also recently been studied within the laboratory; the number of PTB proteins that can be bound has been investigated (Cherny et al., 2010). In this system, weakening the BPS led to a loss of exon three inclusion (Gooding et al., 2006), and this system could be used to investigate the number of U2AF65 bound to constructs with either the wild-type or the weakened branch point sequences.

To study the numbers of U2AF65 and U2AF35 attached to pre-mRNAs, the fluorescent fusion proteins mEGFP-U2AF65, mEGFP-U2AF35, mCherry-U2AF65 and mCherry-U2AF35 were used. These fluorescent fusion proteins were created by cloning the open reading frame of either U2AF65 or U2AF35 into a vector containing the mEGFP or mCherry fluorophores. The N-terminus of U2AF65 or U2AF35 was fused with the C-terminus of mEGFP or mCherry.

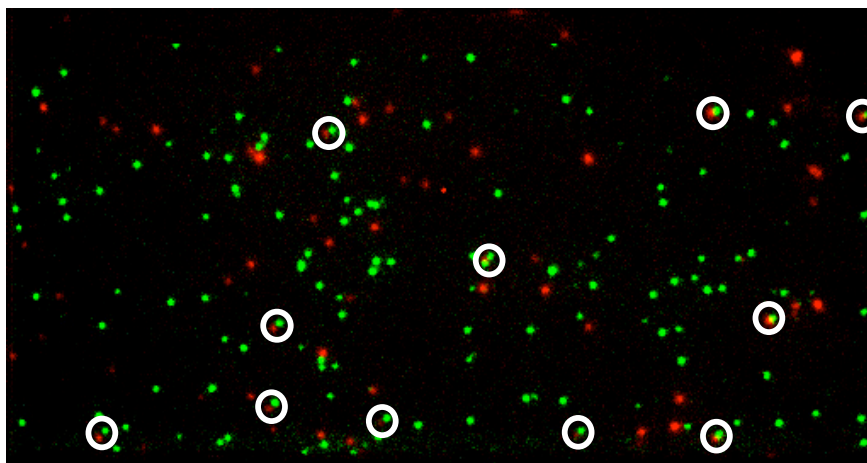
5.1. Preliminary detection of co-localization between mEGFP-U2AF65 and a globin pre-mRNA

As U2AF65 is capable of binding a wide range of sequences, it was feared that the interaction between mEGFP-U2AF65 and a pY tract might be too weak to survive the dilution steps necessary for analysis by TIRF microscopy. A preliminary experiment was carried out whereby HeLa NE containing mEGFP-U2AF65 was depleted of ATP by incubation at 30 °C for 30 minutes and then added to a single molecule splicing reaction, which lacked ATP and Phosphocreatine, but contained the C174C pre-mRNA. A globin construct was chosen because the globin intron contains a short but strong pY tract with the sequence TTCTTCTTTTC and should therefore bind U2AF65. The particular variant, C174C, was chosen because it was readily available.

When the data was collected and analysed, it was clear that co-localization could be observed (Figure 5.1A). To assess the number of U2AF65 proteins that could bind the C174C pre-mRNA, the intensity profiles of mEGFP-U2AF65 spots co-localized with the pre-mRNA were analysed. This analysis proved intriguing, as it clearly showed there was often more than one U2AF65 bound per pre-mRNA; the analysis showed that of the spots that could be assigned, 30 % had a two step intensity profile (Figure 5.1B). The exact number bound could not be calculated, because at this point there was no means of assessing the relative amounts of U2AF65 and mEGFP-U2AF65. However, this preliminary result did suggest that more than one U2AF65 could be bound to the C174C pre-mRNA.

Two possible explanations were considered for this finding, firstly that there were more than one U2AF65 associated with the pY tract of the globin intron and secondly

A



B

Sample	One step	Two step	Unassigned	% One step	% Two step
C174C –ATP	58	26	28	69	31

Figure 5.1: Example of co-localization data for C174C and mEGFP-U2AF65. **(A)** Example of co-localization between C174C pre-mRNA and mEGFP-U2AF65 in the absence of ATP; white circles indicate the co-localized spots. **(B)** Proportions of co-localized spots with a one step, two steps or unassignable intensity profiles are shown. Percentages of one and two step intensity profiles were calculated without the unassigned spots.

that U2AF65 was able to bind elsewhere on the pre-mRNA as well as the pY tract. The possibility that there were more than one U2AF65 associated with the globin pY tract appeared unlikely because of the short nature of the pY tract. The second explanation was plausible because U2AF65 is known to be able to bind pY tracts other than those situated at the 3' SS (Forch et al., 2003) and in some cases, when U2AF65 has been crosslinked to a pre-mRNA, a significant proportion of the signal has been attributed to regions other than the 3' SS (Guth et al., 1999). These studies raised the possibility that U2AF65 may be able to bind the globin pre-mRNAs at sites other than the pY tract. However, it must be noted that it was possible that the extra U2AF65 associating with the pre-mRNA may be due to overexpression.

It was thought possible that part of the reason for the observed co-localization may have been the particular extract being used. For the experiment shown in Figure 5.1, a NE prepared from HeLa cells expressing mEGFP-U2AF65 was used. It is known that U2AF65 alone will select pY tracts (Singh et al., 1995) but that U2AF65 heterodimerised with U2AF35 will select for pY tracts followed by an AG dinucleotide (Wu et al., 1999). Thus U2AF65 may have been binding pY tracts other than at the 3' SS due to the overexpression of mEGFP-U2AF65 relative to the endogenous levels of U2AF35 within the cell.

In order to investigate this possibility, NE expressing different combinations of the U2AF65 and U2AF35 fluorescent fusion proteins was made. Experiments were carried out whereby mEGFP-U2AF65, mCherry-U2AF35 or mEGFP-U2AF65 and mCherry-U2AF35 were transfected into HeLa cells. Transfection of mCherry-U2AF35 on its own killed most of the cells, with only those cells expressing low levels surviving; co-

expression of mEGFP-U2AF65 and mCherry-U2AF35 led to high levels of cell death in cells where mEGFP-U2AF65 expression was low.

To try and solve this problem, a different cell line, HEK-293T, was used. When mCherry-U2AF35 was transfected into HEK-293T cells, aggregates of fluorescent protein formed within the nucleus of the cells (Figure 5.2A and B). These aggregates were not observed when mEGFP-U2AF65 alone was transfected (Figure 5.2C), suggesting that the aggregation was due to the U2AF35. High resolution images of cells transfected with mEGFP-U2AF35 and mEGFP-U2AF65 were taken using a confocal microscope and are shown in Figure 5.2B. This aggregation proved to be a problem, since when NE was made from HEK-293T cells transfected with mEGFP-U2AF35 and analysed by TIRF microscopy, large aggregates of fluorescent protein were observed; this meant that these extracts could not be used (data not shown).

It was hypothesised that the aggregation may be due to improper folding of the U2AF35. U2AF35 had been shown to be unstructured in the absence of U2AF65 and structured in the presence of a peptide derived from U2AF65, *in vitro* (Kellenberger et al., 2002). It was hypothesised that co-transfection of mCherry-U2AF35 with mEGFP-U2AF65 may cause the mCherry-U2AF35 to fold correctly and thus not aggregate. To test this, HEK-293T cells were co-transfected with both mEGFP-U2AF65 and mCherry-U2AF35; when the cells were observed using a light microscope, it was clear that when co-transfected the mCherry-U2AF35 no longer formed aggregates (Figure 5.2D). This suggested that overexpression of U2AF35 relative to U2AF65 leads to misfolding of U2AF35, as there is not enough U2AF65 to form heterodimers. It is hard to fit these observations with the observation that U2AF35 can form homodimers

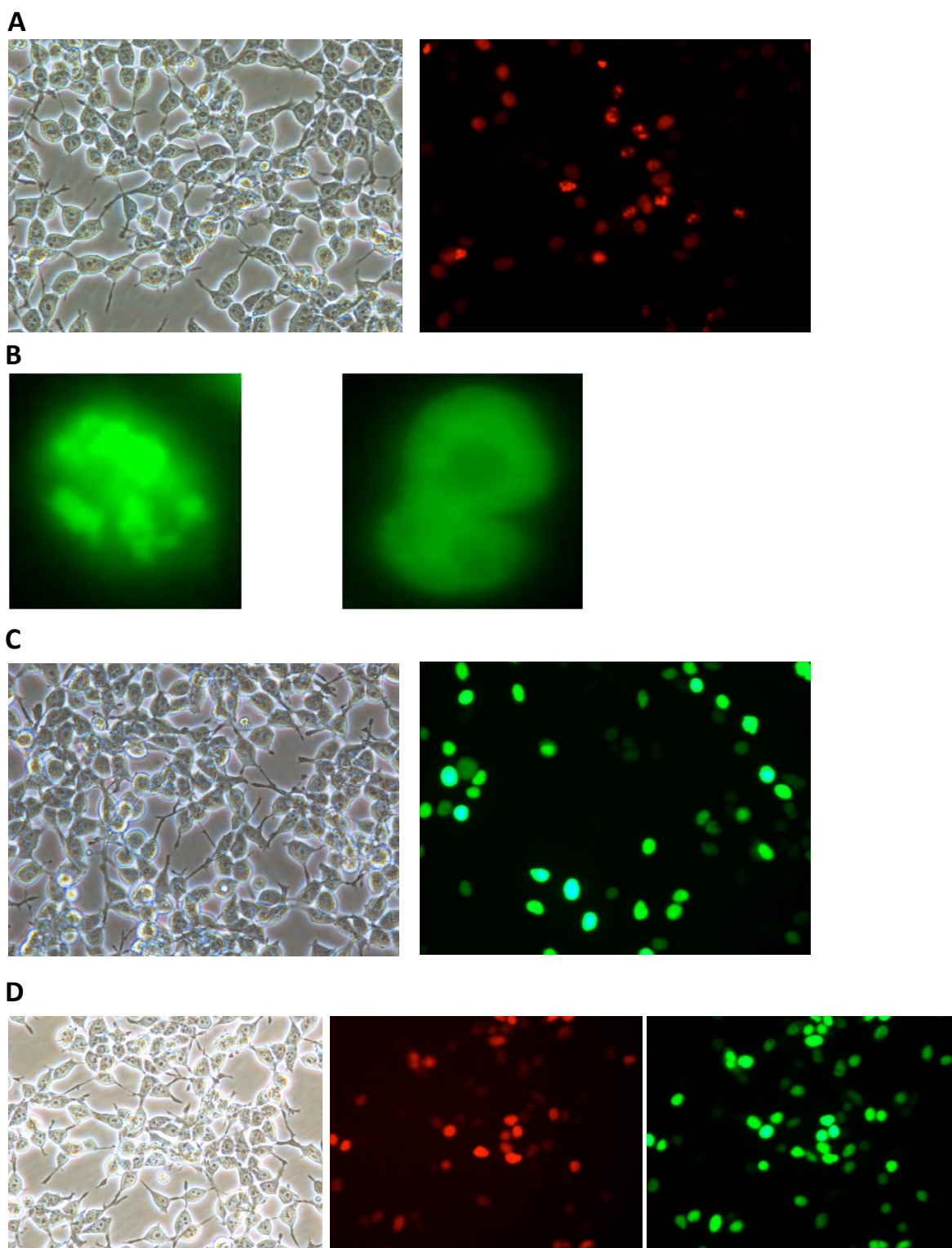


Figure 5.2: Transfection of U2AF65 and U2AF35 constructs into HEK 293T cells. (A) HEK 293T cells transfected with mCherry-U2AF35, the left hand panel shows a bright-field image of the cells, the right hand panel shows an image of the fluorescence under excitation using 633 nm light, imaged using a 40X lens. (B) High magnification (100X) image of structures within the nucleus of HEK 293T cells transfected with mEGFP-U2AF35 (left hand panel) or mEGFP-U2AF65 (right hand panel). (C) HEK 293T cells transfected with mCherry-U2AF65, the left hand panel shows a bright-field image of the cells, panel two is image of fluorescence under excitation using 488 nm light, imaged using a 40X lens. (D) HEK 293T cells co-transfected with mCherry-U2AF35 and mEGFP-U2AF65, the left hand panel is a bright-field image of the transfected cells, the middle panel shows the fluorescence under excitation using 633 nm light and the right hand panel shows the fluorescence under excitation using 488 nm light, imaged using a 40X lens.

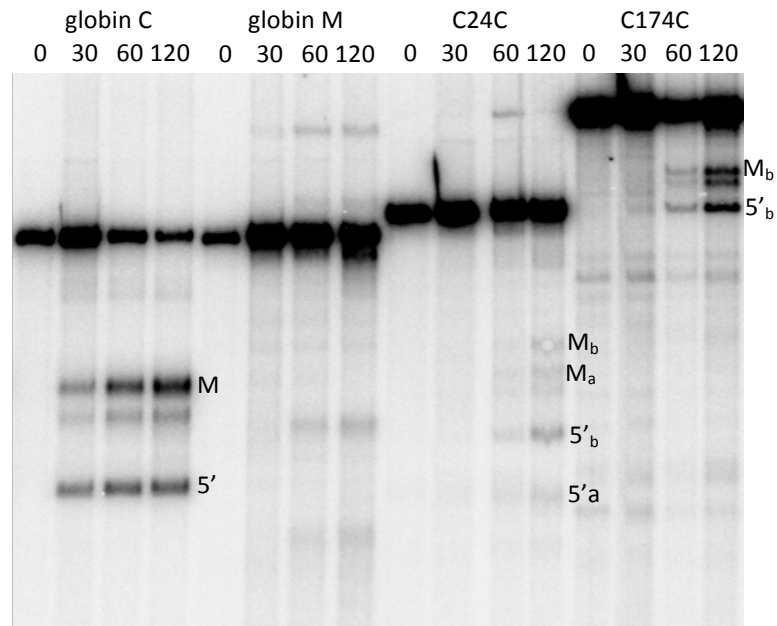
(Chusainow et al., 2005). As a consequence of these results, all later experiments were carried out using NE prepared from HEK-293T cells transfected with equal amounts of plasmid encoding either mEGFP-U2AF65 and mCherry-U2AF35, or mEGFP-U2AF35 and mCherry-U2AF65.

Since the cell line being used had been changed to HEK-293T, it was important to assess whether the pattern of splicing produced by NE made from HEK-293T cells was the same as NE prepared from HeLa cells. The splicing of globin C, globin M, C24C and C174C pre-mRNAs was assessed in standard splicing reactions containing HEK-293T NE. Resolution of the RNA products produced by these reactions on a denaturing urea 6 % polyacrylamide gel showed that the products formed were the same as those produced in HeLa NE (Figures 4.3A and 5.3A). Complex formation was also assessed relative to HeLa NE and no clear difference was observed (Figure 5.3B).

5.2. Functional characterization of U2AF65 and U2AF35 fluorescent fusion proteins

The first step in the characterization of the U2AF65 and U2AF35 fluorescent fusion proteins was to make nuclear extracts from cells expressing them; this allowed the splicing efficiency of the extract to be assessed. Extracts made from HEK-293T cells co-expressing either mEGFP-U2AF65 and mCherry-U2AF35, or mEGFP-U2AF35 and mCherry-U2AF65 spliced efficiently (Figure 5.4A and B). This showed that the fluorescent fusion proteins did not inhibit splicing. Fluorescent Western blots were then carried out on these extracts to assess expression levels. Samples of the two extracts were probed with both the α U2AF35 and α U2AF65 antibodies (Table 2.3). As the source of the two antibodies was different, two secondary antibodies conjugated to two fluorophores (α Mouse and α Rabbit, Table 2.3) could be used for simultaneous

A



B

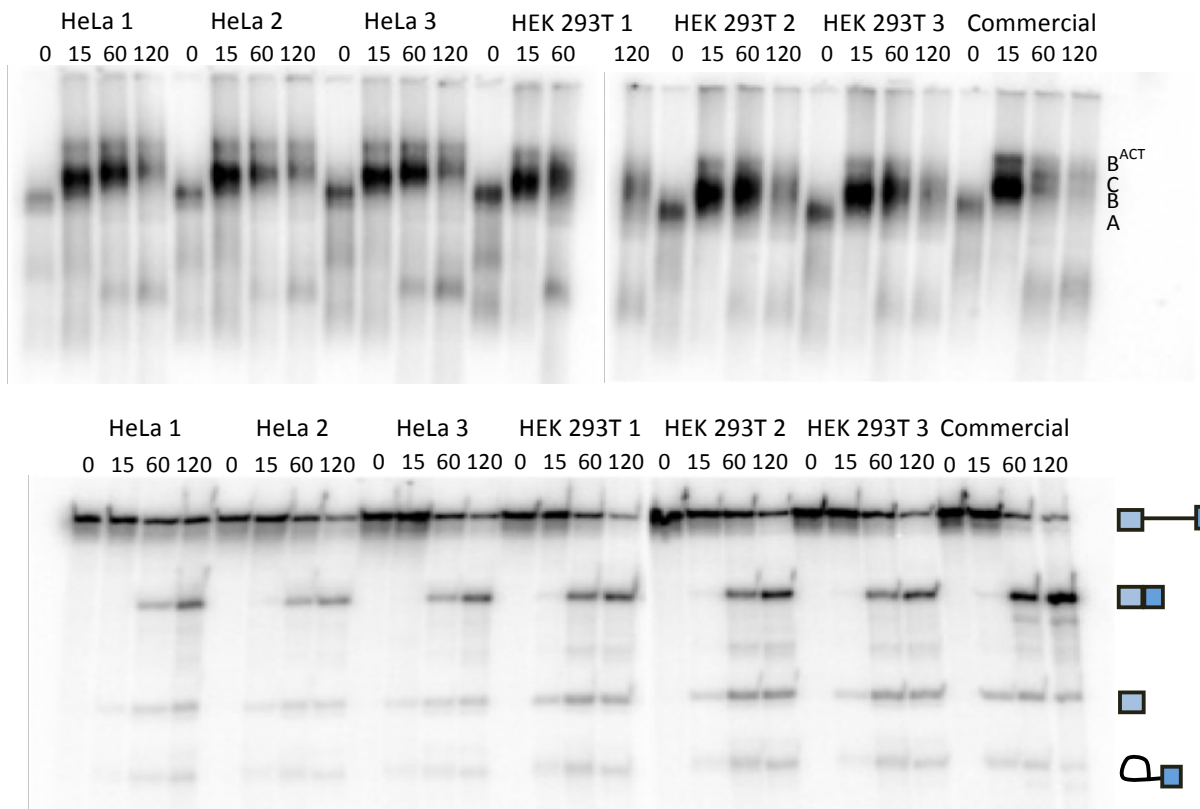


Figure 5.3: Splicing and complex formation in HEK 293T NE. (A) A denaturing urea 6 % polyacrylamide gel showing splicing of globin C, globin M, C24C and C174C pre-mRNAs in HEK 293T NE. **(B)** Complex formation in HEK 293T NE, the top panel shows complex formation in either three HeLa, HEK 293T, or commercial NE. The bottom panel shows the RNA products formed in the splicing reactions used in the top panel run on a denaturing urea 6 % polyacrylamide gel.

detection of the signals from both antibodies. The images of these blots showed that each of the fluorescent proteins was expressed in the NE. Interestingly, both mCherry-U2AF35 and mCherry-U2AF65 constructs produced two bands (Figure 5.4C). When bacterially expressed mCherry is run on a standard denaturing SDS polyacrylamide gel it also produces two bands; when these two bands were cut out and analysed by mass spectroscopy, both were found to be full length, suggesting that mCherry is partially resistant to the denaturing conditions normally used (Dr S. Badyal, personal communication). As the mCherry-U2AF35 and mCherry-U2AF65 were part of a heterogeneous nuclear extract, excision from a gel and mass spectroscopy would not be feasible, so this phenomenon was investigated by another means. A second denaturing SDS polyacrylamide gel was run that contained 7M urea to make the gel more denaturing. Analysis by fluorescent Western blot of the mCherry-fusion proteins run on this highly denaturing gel showed that the intensity of the faster migrating band in both cases was reduced relative to the slower migrating band (Figure 5.4D). Proportional analysis of the signal in each band showed that for mCherry-U2AF35 the ratio between the top band and lower band was reduced from 1:0.77 in the absence of urea to 1:0.28 in the presence of urea; for mCherry-U2AF65 the ratio was reduced from 1:0.6 to 1:0.25 indicating that the two bands are likely to be full-length fusion protein.

The ability of mEGFP-U2AF65 and mEGFP-U2AF35 to bind pre-mRNA was assessed by immunoprecipitation of globin C pre-mRNA from splicing reactions containing the two fusion proteins, using the α EGFP antibody (Table 2.3). globin C pre-mRNA was efficiently immunoprecipitated from splicing reactions containing either mEGFP-

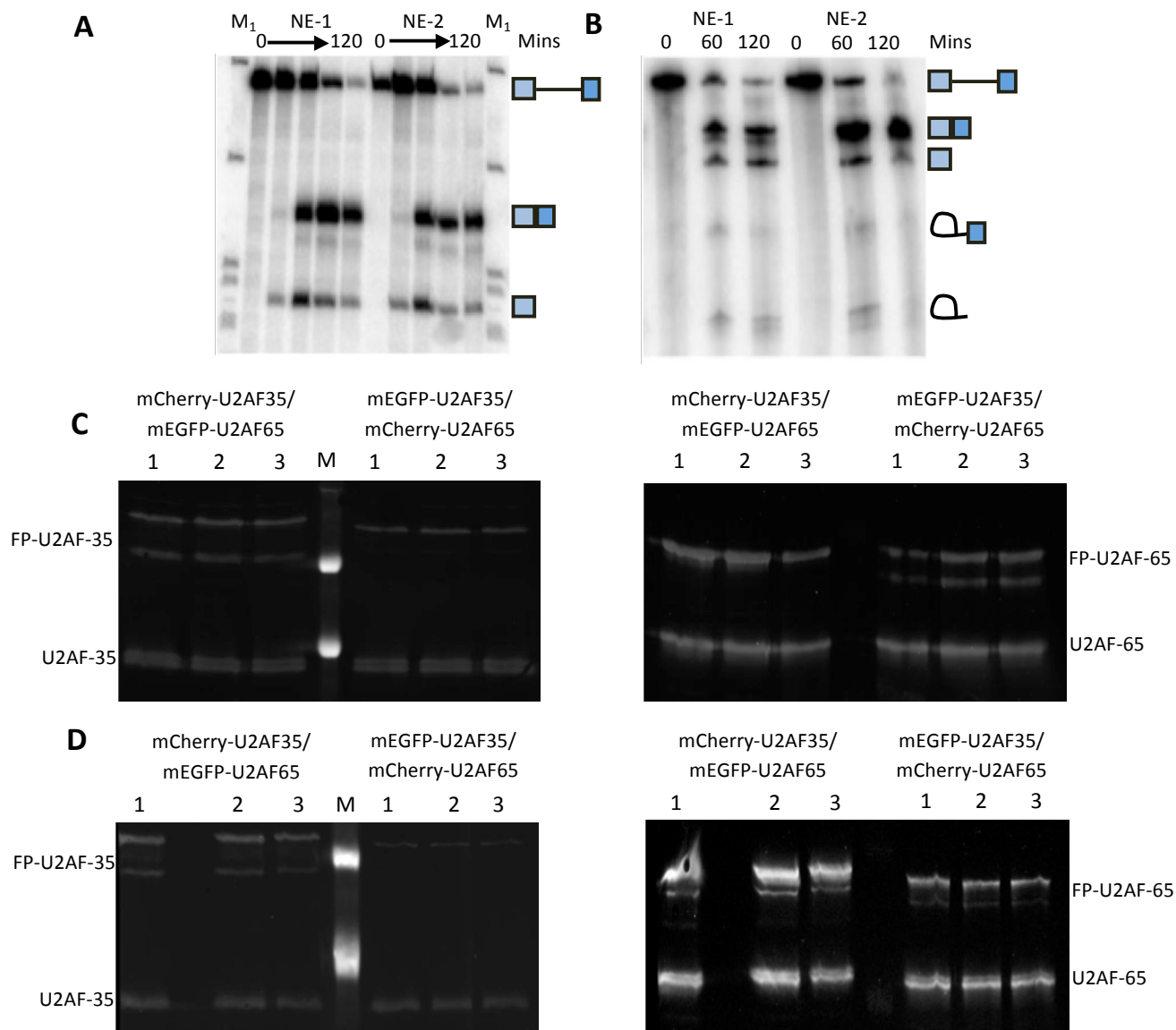


Figure 5.4: Characterization of NE. **(A)** RNA products of a splicing reaction run on a denaturing urea 6 % polyacrylamide gel, which was carried out using globin C pre-mRNA and HEK-293T NE containing mEGFP-U2AF65 and mCherry-U2AF35 (NE-1), or commercial NE (NE-2); timepoints of 0, 15, 30, 60 and 120 minutes were taken. **(B)** RNA products of a splicing reaction run on a denaturing urea 6 % polyacrylamide gel, which was carried out using globin C pre-mRNA and HEK-293T NE containing mCherry-U2AF65 and mEGFP-U2AF35 (NE-1), or commercial NE (NE-2). **(C)** Fluorescent western blot of NE containing either mCherry-U2AF35 and mEGFP-U2AF65, or mCherry-U2AF65 and mEGFP-U2AF35. The left hand panel shows a Western blot using the α U2AF35 antibody, and the right hand panel shows a Western blot using the α U2AF65 antibody (Table 2.3). **(D)** A Western blot performed as in Figure 5.4C except that the gel contained 7M urea.

U2AF65 or mEGFP-U2AF35 showing association of both with the pre-mRNA (Figure 4.1E and Figure 5.5A). Globin C pre-mRNA was also efficiently immunoprecipitated in the presence of ATP from splicing reactions containing mEGFP-U2AF35 (Figure 5.5A). Immunoprecipitation of the mCherry versions of U2AF65 and U2AF35 was not carried out due to a lack of suitable antibodies.

As a second test of the ability of mEGFP-U2AF65 to associate with a pre-mRNA, mEGFP-U2AF65 was cross-linked to the α -tropomyosin exon 3 3' SS region using ultra-violet cross-linking. The fragment of pre-mRNA used for this experiment was transcribed from a plasmid containing the TM1 fragment of rat α -tropomyosin (Cherny et al., 2010) that had been digested using PvuII, so that the transcript would run from the BPS to the beginning of exon three. The transcription was carried out in the presence of [α -³²P] UTP, in order to enhance the cross-linking signal because U2AF65 will protect runs of uridine bases. Cross-linking of the mEGFP-U2AF65 fluorescent fusion protein to this fragment of pre-mRNA was clearly detected (Figure 5.5B); the band could be identified due to its size and due to its absence in HEK 293T NE made from untransfected cells. Endogenous U2AF65 was identified based on its size and its dependence on KCl concentration: cross-linking of U2AF65 to the α -tropomyosin exon 3 3' SS is dependent on the concentration of KCl in the splicing reaction; as the KCl concentration is increased, U2AF65 cross-linking becomes detectable, whilst PTB cross-linking is reduced (Dr C.Gooding, personal communication; Figure 5.5C).

The two subunits of U2AF, which are known to interact with each other through a proline sandwich (Kielkopf et al., 2001), should be able to interact with each other if the fluorescent fusion proteins are functional. To test this, TIRF microscopy was used.

NE containing mEGFP-U2AF65 and mCherry-U2AF35 was treated with RNase A for 15 minutes at 30 °C at a concentration of 1.4 µg/µl to degrade the RNA in the NE. The extract was then diluted, ready for analysis by TIRF microscopy. As biotinylated oligonucleotide was not being used, a different method was needed to attach the sample to the surface. 30 µl of the αDSRed antibody (Table 2.3) diluted to a concentration of 1:500 in dilution buffer (Section 2.5) was injected into the flow-cell and incubated at room temperature in order to non-specifically absorb onto the surface of the silica slide. The slide was then washed with dilution buffer before the diluted NE was added to the flow cell. When the fluorescence signals produced by mCherry-U2AF35 and mEGFP-U2AF65 were overlayed it was clear there was significant levels of co-localization (Figure 5.5D), showing that the interaction was RNA independent.

Overall it was concluded that the fluorescent fusion proteins used for TIRF analysis were functional. No experiments were carried out directly on the mCherry-U2AF65 fusion protein, but the experiments showing that mEGFP-U2AF65 is functional and the observation that mCherry-U2AF65 stopped aggregates of mEGFP-U2AF35 forming in the nuclei of HEK-293T cells support that it was functional.

5.3. Co-localization between mEGFP-U2AF65 and a 'negative control' pre-mRNA

As described in Section 5.1, one possible reason for there being more than one U2AF65 bound to the C174C pre-mRNA was that U2AF65 is able to bind in the absence of U2AF35 and so binds at positions it would not normally. To assess whether co-transfection of mCherry-U2AF35 with mEGFP-U2AF65 led to a loss of non-specific binding a series of experiments were carried out; in the first experiment, as a control,

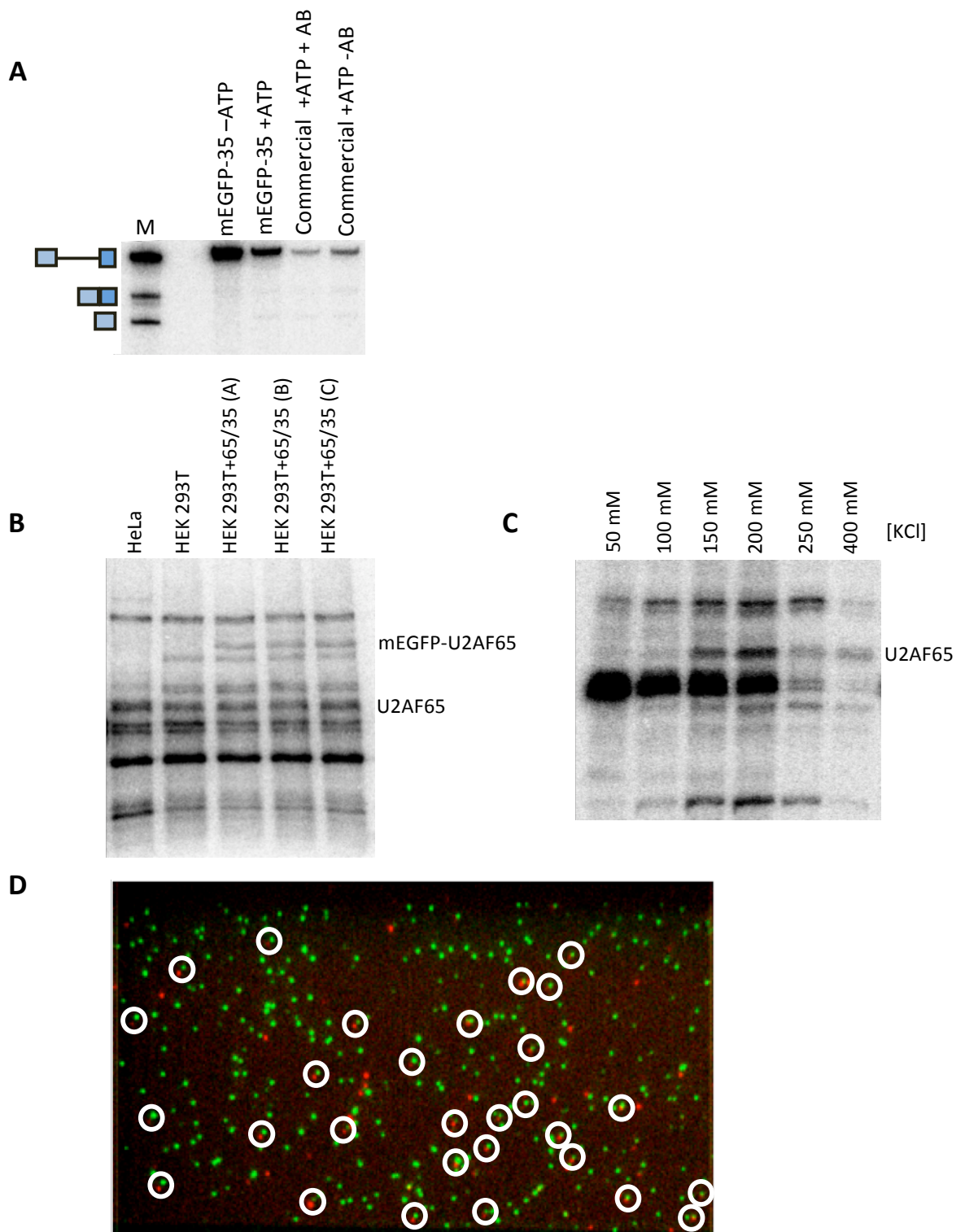


Figure 5.5: Functional assays for mEGFP-U2AF35 and mEGFP-U2AF65. **(A)** Denaturing urea 6 % polyacrylamide gel showing immunoprecipitation of pre-mRNA bound to mEGFP-U2AF35, in the absence or presence of ATP. Immunoprecipitation was performed using the α EGFP antibody and NE co-expressing mEGFP-U2AF35 and mCherry-U2AF65, or commercial NE. Commercial NE immunoprecipitations were carried out with (+AB) or without (-AB) antibody. **(B)** Denaturing SDS 10 % polyacrylamide gel showing cross-linking of proteins to radiolabelled tropomyosin exon 3-pY tract, cross-linked in splicing reactions containing 250 mM KCl. Cross-linking was carried out using NE made from, HeLa, HEK-293T, or HEK-293T transfected with mEGFP-U2AF-65 and mCherry-U2AF-35. **(C)** Denaturing 10 % SDS polyacrylamide gel showing crosslinking of proteins to radiolabelled tropomyosin exon 3-pY tract in splicing reactions containing the indicated concentration of KCl. **(D)** Example image showing co-localization between mEGFP-U2AF65 and mCherry-U2AF35 in NE pre-treated with RNase A.

co-localization between globin C and mEGFP-U2AF65 was assessed when an extract co-expressing mEGFP-U2AF65 and mCherry-U2AF35 was used. The globin C construct was labelled as described in Section 4.2 using the Bg-5'-Cy5 oligonucleotide (Table 4.2). For this experiment, single molecule splicing reactions were set up lacking ATP (E complex conditions) as described in section 2.3.4. Analysis of the images taken shows that, as expected, co-localization was still observed (Figure 5.6A).

To assess whether this co-localization was dependent upon the pY tract of globin C, a new construct was made in which the pY tract of globin C was deleted. Deletion of the globin C pY tract was achieved by amplification of two fragments from the globin C construct by PCR. The first fragment was amplified using the P16 (Table 2.1) and Bg pY del 1 (Table 5.1) oligonucleotides; this fragment extended from the 5' terminus of the transcript to the BPS within the globin C intron. The second fragment was amplified using the Bg pY del 2 (Table 5.1) and β G E3 no 5' SS (Table 2.1) oligonucleotides; this fragment extended from the 3' SS of globin C to the 3' terminus of the construct. The Bg pY del 1 oligonucleotide was designed to have a region of 20nt overlap with the 5' end of the second fragment, so when fragment one and two were added to a single PCR reaction containing the P16 and β G E3 no 5' SS oligonucleotides a globin C variant lacking the pY tract was produced.

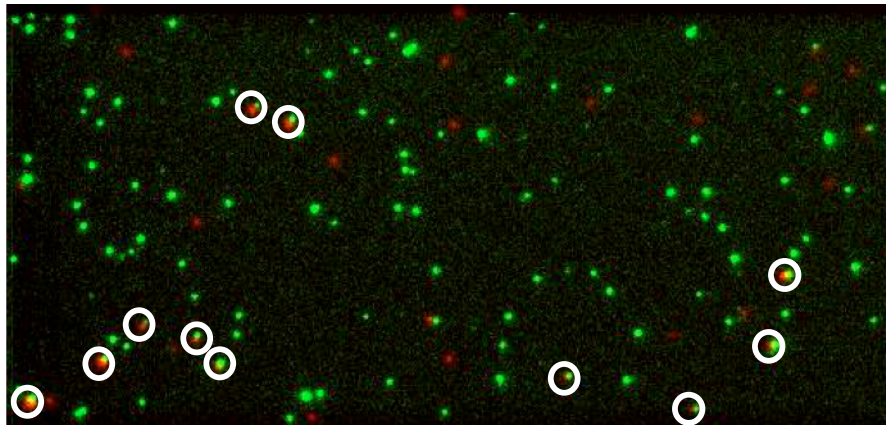
The construct, globin C Δ pY, did not splice in a standard splicing reaction (Figure 5.6C). However, despite not having a pY tract, co-localization with mEGFP-U2AF65 was still observed (Figure 5.6B), this experiment was carried out under E complex conditions. This result confirmed that mEGFP-U2AF65 was able to associate with the globin C pre-mRNA at positions other than the pY tract used in splicing. One possible explanation

for the observed co-localization is that it is mediated by the U1 snRNP bound to the 5' SS, since there have been reports that U1 snRNP influences the binding of U2AF65 to the 3' SS both upstream (Hoffman and Grabowski, 1992) and downstream (Li and Blencowe, 1999; Cote et al., 1995) of the 5' SS. U1 snRNP in conjunction with SF1 has also been suggested to be important in U2AF65 recruitment, based on the characterization of an E like complex, E'. (Kent et al., 2005). To assess whether recruitment of U2AF65 to the globin C pre-mRNA was dependent on the presence of a U1 snRNP bound to the 5' SS, an experiment was carried out to determine whether mEGFP-U2AF65 binds globin M. Globin M lacks a 5' SS and so cannot bind U1 snRNP (Section 4.4); therefore, if co-localization of mEGFP-U2AF65 with globin C is dependent upon U1 snRNP, no co-localization should have been detected. Analysis of the images obtained from an experiment carried out in the absence of ATP showed that there was significant co-localization (Figure 5.6D). This result showed that mEGFP-U2AF65 co-localization with globin C, containing a pY tract, is not dependent on U1 snRNP being bound to the 5' SS.

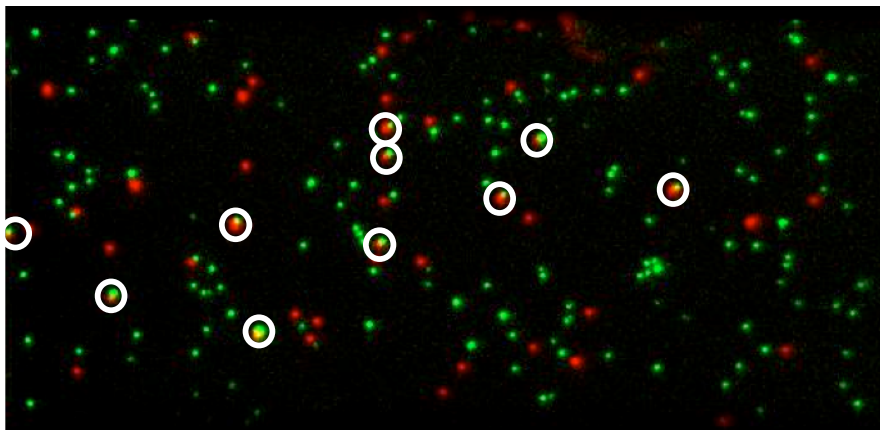
Name	Sequence
Bg pY del 1	CAGCACGTTGCCAGGAGCTGTTAGCAGAGGGGCCCCGGTT
Bg pY del 2	AGCTCCTGGGCAACGTGCTG
Bg Cryptic pY 1	CATGAATTTTACAATAGCGCAATCAAGGGTCCCCAACT
Bg Cryptic pY 2	CGCTATTGTAAAAATTCATGTTATATGGTC
gloC cpY mut F	TTTGGGGACCCTTGATTGTTTCATTGATATGCGCTATTGTAAAAATTCATG
glo C cpy mut R	CATGAATTTTACAATAGCGCATATCAATGAACAATCAAGGGTCCCCAAA
X-pY-trx F	AAATTAATACGACTCACTATAGGCCTAACCATGTTTCATGCCTT
X-pY R	TGCCAGGAGCTGTAGG

Table 5.1: Oligonucleotides used for the construction of pre-mRNA templates.

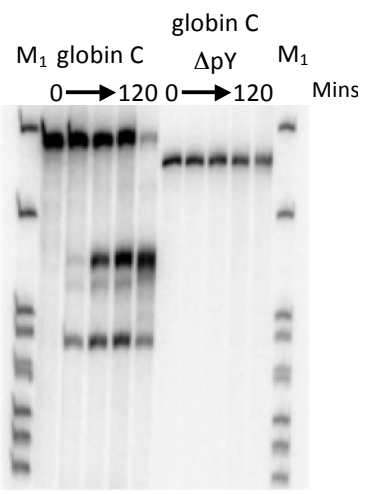
A



B



C



D

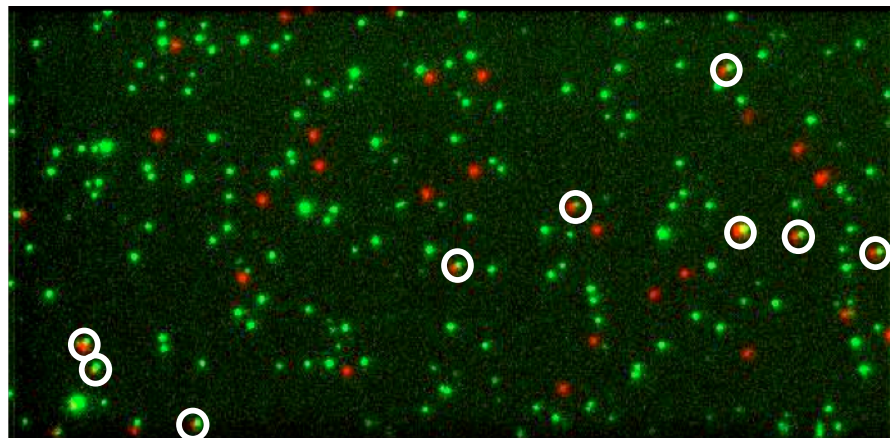


Figure 5.6: Example co-localization data for mEGFP-U2AF65. **(A)** Example of co-localization between Cy5 oligonucleotide-labelled globin C pre-mRNA and mEGFP-U2AF65 in the absence of ATP, white circles indicate the co-localized spots. **(B)** Example of co-localization between globin C Δ pY and mEGFP-U2AF65 in the absence of ATP. **(C)** Denaturing urea 6 % polyacrylamide gel of the products formed by splicing globin C or globin C Δ pY in NE. **(D)** Example of co-localization between globin M and mEGFP-U2AF65 in the absence of ATP.

5.4. Analysis of the 24-36U/C pY tract of globin C

To determine how U2AF65 may be associating with the globin C Δ pY pre-mRNA the globin C sequence was studied. Analysis of the globin C sequence showed that there was a second strong pY tract within the intron of the construct. This second pY tract was 24nt downstream of the 5' SS and had almost the same sequence as the pY tract needed for splicing (Figure 5.7A). As this second pY tract, known as 24-36U/C, was very close in sequence to the pY tract used in splicing it seemed possible that mEGFP-U2AF65 was binding to it and thus causing co-localization to be detected in the absence of the normal pY tract.

To investigate the function of this 24-36U/C region of globin C two deletion mutants were constructed. First, globin C Δ 24-36U/C was made in which the 24-36 U/C region was removed. This was done using essentially the same method used to create the globin C Δ pY in section 5.3, except that the Bg pY del1 oligonucleotide was replaced with the Bg Cryptic pY 1 oligonucleotide and the Bg pY del 2 oligonucleotide was replaced with the Bg Cryptic pY 2 oligonucleotide (Table 5.1). The second construct was the globin C Δ pY, Δ 24-36U/C in which the 24-36U/C region was removed from a construct already lacking the pY tract used for splicing.

In order to assess the effect of these deletions on splicing efficiency the globin C, globin C Δ 24-36U/C, globin C Δ pY and globin C Δ pY, Δ 24-36U/C pre-mRNAs were spliced in standard splicing reactions. Deletion of the 24-36U/C led to a considerable decrease in splicing efficiency (Figure 5.7B). When the splicing efficiencies of globin C and globin C Δ 24-36U/C were quantified, by calculating the percentage of the total counts that each RNA species represented, it was clear that after 60 minutes incubation at 30 °C

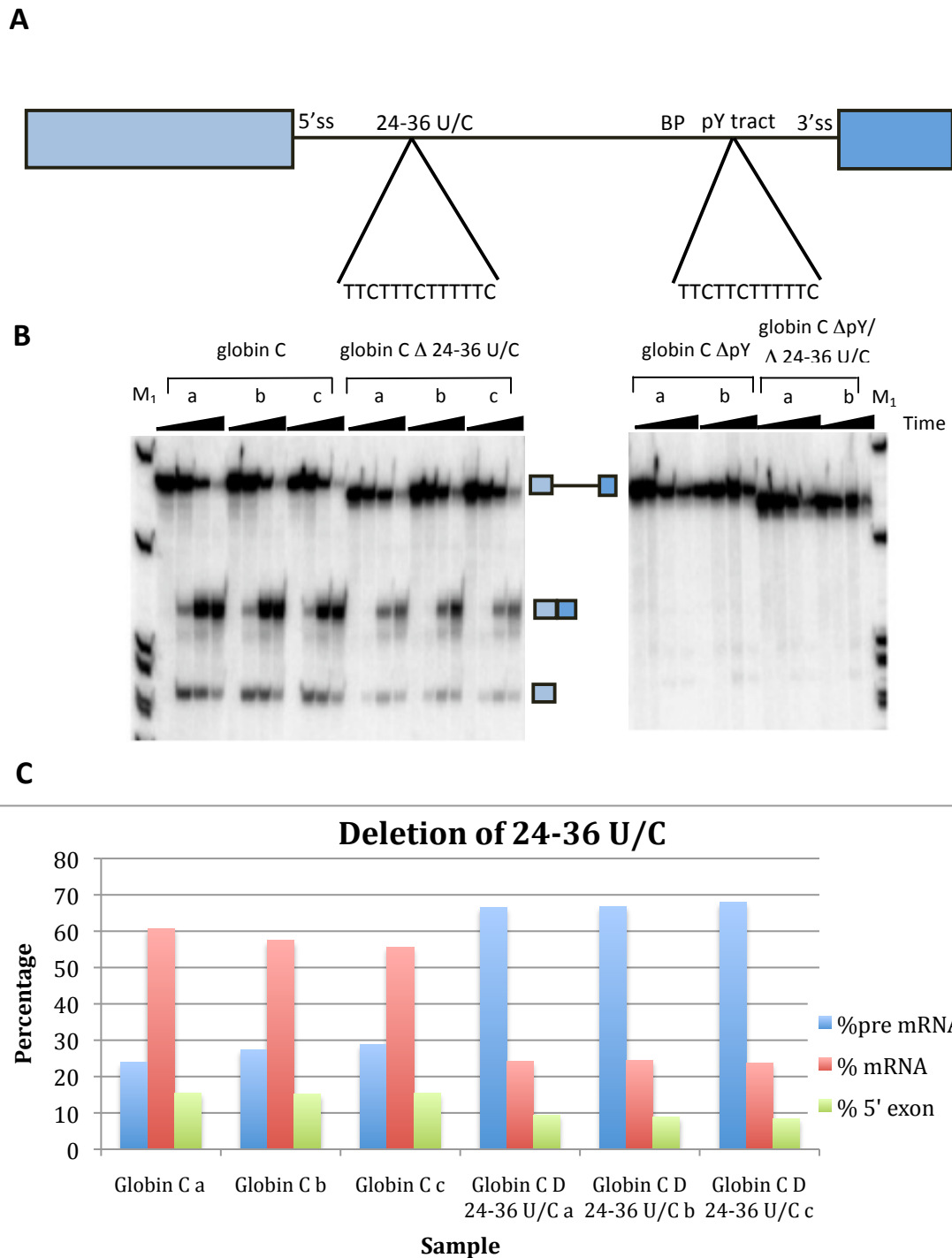


Figure 5.7: Deletion of the 24-36 U/C region. (A) Diagrammatic representation of the approximate position of the 24-36 U/C region within globin C. (B) A denaturing urea polyacrylamide gel showing the effect of deletion of 24-36 U/C on the efficiency of splicing of globin C and globin C Δ pY, timepoints were taken at 0, 30, 60 and 120 minutes. (C) Quantification of globin C and globin C Δ 24-36 U/C splicing. The 60-minute time point was used for quantification and data was corrected for the proportion of guanosine nucleotides in each product. Samples labelled globin C D 24-36 U/C are the same as those labelled globin C Δ 24-36 U/C.

globin C $\Delta 24-36\text{U/C}$ (~22 % spliced) was severely inhibited relative to globin C (~55 % spliced; Figure 5.7C). The deletion of 24-36 U/C from the construct lacking the pY tract at the 3' SS (globin C Δ pY, Δ 24-36U/C) did not affect splicing as the globin C Δ pY pre-mRNA is unable to splice (Figure 5.7B).

This finding suggested that 24-36 U/C was crucial for efficient splicing of the globin C pre-mRNA. Although U2AF65 had been shown to bind pY tracts close to the 5' SS (Forch et al., 2003), other proteins had also been shown to bind similar sequences, such as Tia1 (Forch et al., 2002) and PTB (Shukla et al., 2005).

To test whether deletion of the 24-36U/C region may have been inhibitory due to the loss of Tia1 binding, the level of complex formation on constructs containing only the 5' portion of the globin C intron was assessed. This assay was used because Tia 1 enhances the binding of U1 snRNP to the 5' SS through the U1 snRNP protein U1C (Forch et al., 2002) and so U1 snRNP dependent complexes should be inhibited if Tia 1 is important, as has been observed previously (Forch et al., 2002). Pre-mRNAs lacking all sequence downstream of the BPS ($\Delta 3'$) were used for this study. Analysis of complex formation on the globin C $\Delta 3'$, globin M $\Delta 3'$ and globin C $\Delta 24-36\text{U}/\Delta 3'$ premRNAs showed that whilst globin M $\Delta 3'$ formed only H complex, both globin C $\Delta 3'$ and globin C $\Delta 24-36\text{U}/\Delta 3'$ could form larger complexes. These complexes were observed both in the presence and absence of ATP (Figure 5.8A). This made it unlikely that Tia 1 was causing the effect, because in previous reports inhibition of Tia 1 binding led to complete complex loss (Forch et al., 2002).

To investigate the point at which splicing was being inhibited, spliceosomal complex formation on the globin C and globin C $\Delta 24-36\text{U/C}$ pre-mRNAs was compared.

Analysis of complex formation over a 30 minute time course showed that the B to B^{ACT} complex conversion was inhibited by loss of the 24-36U/C region (Figure 5.8B). This result was unexpected as any protein known to bind similar pY tracts effects early steps in spliceosomal assembly (Forch et al., 2002; Forch et al., 2003; Shukla et al., 2005). However, this experiment showed that the deletion was affecting later stages of spliceosomal assembly, possibly affecting rearrangements necessary for the B to B^{ACT} complex conversion.

To assess whether mEGFP-U2AF65 can bind the 24-36U/C region of globin C, a competitive Immunoprecipitation was performed. In this, three RNAs of different lengths were added to a splicing reaction made using NE containing mEGFP-U2AF65 and mCherry-U2AF35. The first RNA was a small fragment, 45nt long, called the X-pY fragment, which comprised the BPS, pY tract and 3' SS AG dinucleotide of the globin C pre-mRNA. It was transcribed from a PCR product that was made using the X-pY-trxF and X-pYR oligonucleotides (Table 5.1). The second and third RNAs used the globin CΔ3' and globin C Δ24-36U/CΔ3', these lacked all sequence downstream of the BPS. Therefore, U2AF65 was expected to bind X-pY well, and if it bound the 24-36U/C efficiently then globin C Δ3' was expected to be selected relative to globin C Δ24-36U/CΔ3'. These three RNAs were added to a splicing reaction, lacking ATP, and an immunoprecipitation of the RNA bound to mEGFP-U2AF65 was carried out using the αEGFP antibody (Table 2.3). The RNAs were purified and run on a dis-continuous denaturing urea 6-15 % polyacrylamide gel. A discontinuous gel was used to allow resolution of all three bands on a single gel (Figure 5.8C). To assess which RNAs were precipitated most efficiently the relative proportions of each RNA added to the splicing

reaction were calculated from the input samples and then compared to the relative proportions of each RNA that had been immunoprecipitated. This is displayed as percentage change of total in Figure 5.8C. These results show that the X-pY RNA was preferred compared with the globin C $\Delta 3'$ and globin C $\Delta 24-36\text{U}/\text{C}\Delta 3'$ RNAs. However, importantly, the globin C $\Delta 24-36\text{U}/\text{C}\Delta 3'$ RNA was selected against to a greater extent than the globin C $\Delta 3'$ suggesting that removal of the 24-36U/C region had reduced the affinity of the RNA towards mEGFP-U2AF65. As can be observed in Figure 5.8 the X-pY RNA degraded considerably during the immunoprecipitation, this was hypothesised to be because the X-pY RNA had been transcribed in the absence of diguanosine triphosphate cap, leading to degradation upon addition to a splicing reaction. For the quantification all degradation bands were included in the X-pY sample.

The effect of deleting the 24-36U/C pY tract on co-localization between mEGFP-U2AF65 and the globin C pre-mRNA, at the single molecule level, was assessed using the globin C $\Delta \text{pY}, \Delta 24-36\text{U}/\text{C}$ pre-mRNA as this had neither of the two identified pY tracts. Therefore if the observed co-localization between globin C ΔpY and mEGFP-U2AF65 was due to the 24-36U/C pY tract, there should be no co-localization between globin C $\Delta \text{pY}, \Delta 24-36\text{U}/\text{C}$ and mEGFP-U2AF65.

The globin C $\Delta \text{pY}, \Delta 24-36\text{U}/\text{C}$ pre-mRNA was first analysed under E complex conditions. For this, the NE containing mEGFP-U2AF65 and mCherry-U2AF35 was pre-incubated at 30 °C for 30 minutes to deplete ATP, before addition to a single molecule splicing reaction, lacking ATP and Phosphocreatine and analysis by TIRF microscopy. Co-localization was still observed showing that the 24-36U/C pY tract was not the cause of the co-localisation being observed (Figure 5.9A). To try and understand where the co-

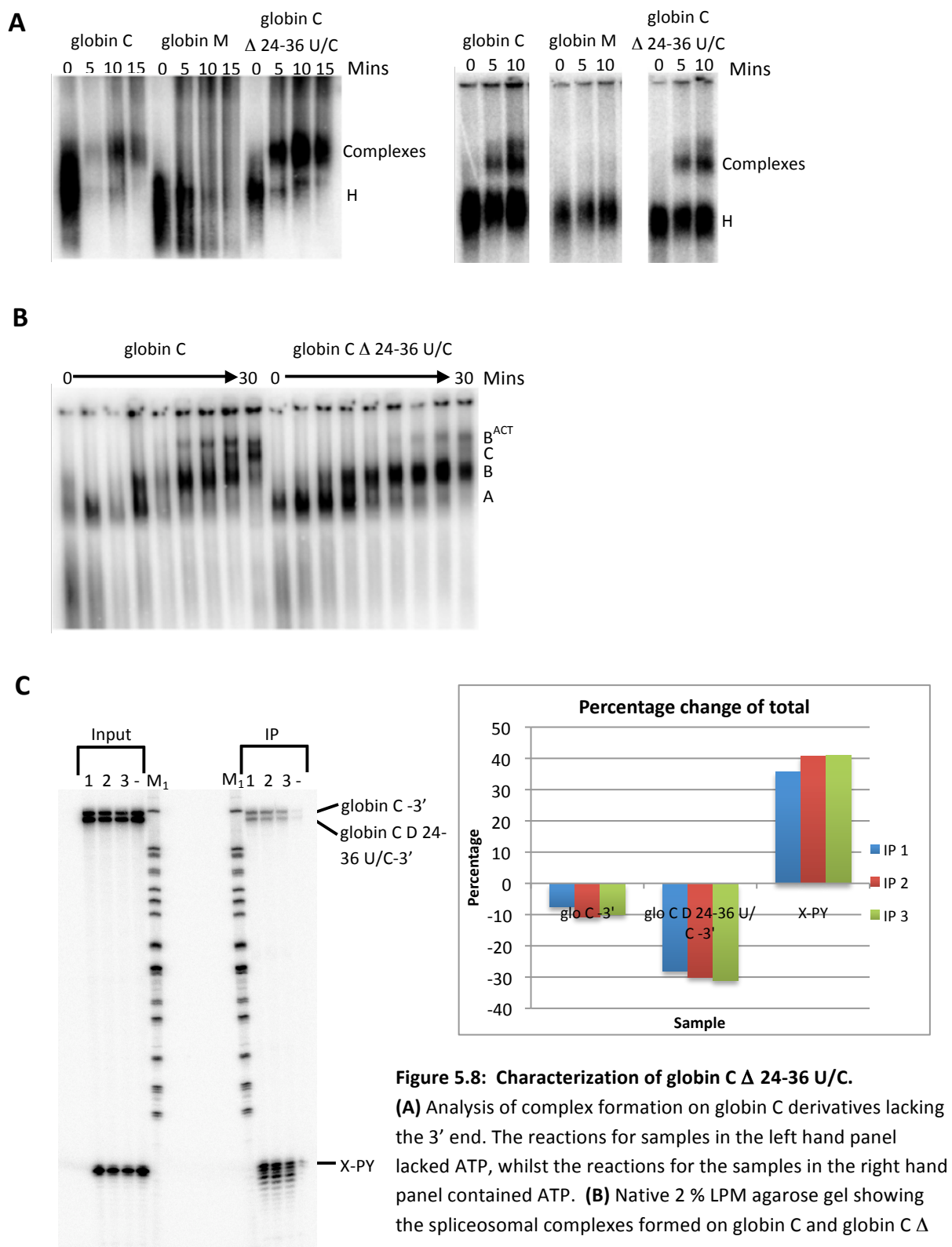


Figure 5.8: Characterization of globin C Δ 24-36 U/C.

(A) Analysis of complex formation on globin C derivatives lacking the 3' end. The reactions for samples in the left hand panel lacked ATP, whilst the reactions for the samples in the right hand panel contained ATP. **(B)** Native 2 % LPM agarose gel showing the spliceosomal complexes formed on globin C and globin C Δ 24-36 U/C. Time points were taken at 0, 1, 2.5, 5, 7.5, 10, 12.5, 15 and 30 minutes. **(C)** Competitive immunoprecipitation performed using HEK 293T NE containing mEGFP-U2AF65 and mCherry-U2AF35 and the αEGFP antibody. The left hand panel shows the 1 % input and immunoprecipitated RNAs run on a denaturing urea 6-15 % discontinuous polyacrylamide gel, the right hand panel shows the change in the relative proportions of each RNA for each immunoprecipitation.

localization was coming from a series of different conditions were examined. Firstly an experiment was carried out in which, if a splicing capable pre-mRNA was used, complex assembly would be stalled at A complex. This was by pre-incubation of the NE with 1 μ M α -U6 oligonucleotide (Table 2.1) at 30 $^{\circ}$ C for 15 minutes prior to addition to a splicing reaction containing ATP and Phosphocreatine. Analysis of the images obtained showed co-localization was still observed (Figure 5.9B). The second condition used was designed to cause sequestration of U1 snRNP under E complex conditions. To achieve this α -U1 oligonucleotide was added to the NE, at a concentration of 4 μ M, whilst it was being pre-incubated at 30 $^{\circ}$ C for 30 minutes to deplete ATP. Again under these conditions co-localization was still observed (Figure 5.9C). This suggests that in the absence of ATP, co-localization between globin C Δ pY, Δ 24-36U/C and mEGFP-U2AF65 is not dependent on the presence of U1 snRNP. However, under the third condition tested, in which the NE was pre-incubated at 30 $^{\circ}$ C for 15 minutes in the presence of 1 μ M α -U6 oligonucleotide and 4 μ M α -U1 oligonucleotide, co-localization dropped from 20 % seen under A complex conditions to 4 % (Figure 5.9D; Table 5.2). Hence, U1 snRNP played a major role in the recruitment of U2AF65, to sites other than the pY tract used in splicing, in an ATP-dependent manner.

These results from the TIRF microscope led to the conclusion that the 24-36U/C region was not the cause of the observed co-localization between globin C Δ pY and mEGFP-U2AF65. Instead they raised the interesting possibility that U1 snRNP may influence binding of U2AF65 to globin C pre-mRNA. The effect of deletion of the 24-36U/C region on splicing efficiency and complex formation is thought to be due to a

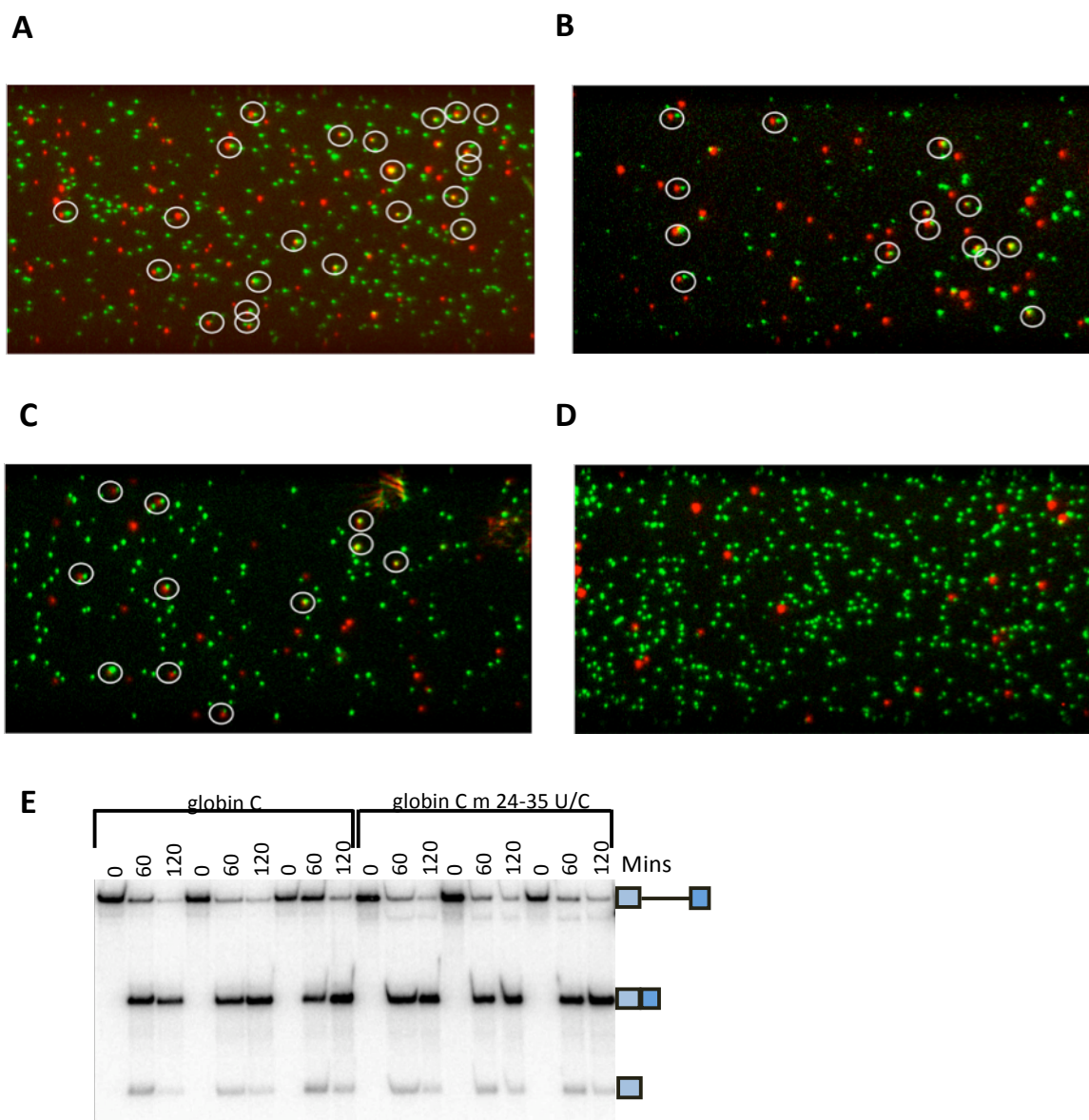


Figure 5.9: Co-localization between globin C $\Delta pY, \Delta 24-36$ U/C and mEGFP-U2AF65. (A) Image showing the co-localization between globin C $\Delta pY, \Delta 24-36$ U/C and mEGFP-U2AF65 in the absence of ATP. (B) Image showing the co-localization between globin C $\Delta pY, \Delta 24-36$ U/C and mEGFP-U2AF65 in the presence of ATP. (C) Image showing the co-localization between globin C $\Delta pY, \Delta 24-36$ U/C and mEGFP-U2AF65 in the absence of ATP, with the addition of the α -U1 oligonucleotide (Table 2.1). (D) Image showing no co-localization between globin C $\Delta pY, \Delta 24-36$ U/C and mEGFP-U2AF65 in the presence of ATP and the α -U1 oligonucleotide. (E) Splicing products formed after incubation of globin C or globin C m 24-36 U/C in HeLa

Sample	Total Cy5 spots	Co-localized Cy5 spots	Percentage co-localization
globin C Δ pY, Δ 24-36 U/C -ATP	567	118	21
globin C Δ pY, Δ 24-36 U/C +U6	353	93	26
globin C Δ pY, Δ 24-36 U/C –ATP +U1	50	12	20
globin C Δ pY, Δ 24-36 U/C +U6 +U1	213	10	4

Table 5.2: Efficiency of co-localization between globin C Δ pY, Δ 24-36 U/C and mEGFP-U2AF65.

shortening of the intron as mutation of the 24-36U/C sequence to TTCATTGATATGC did not lead to any effect on splicing efficiency (Figure 5.9E). The globin C m24-36U/C construct was created by mutagenesis of the globin C plasmid using the globin C cpy mut F and globin C cpy mut R oligonucleotides (Table 5.1). In the globin C Δ 24-36U/C the deletion had caused a shortening of the intron from 107nt to 93nt, which may have caused the inhibition of splicing and complex assembly (Figure 5.7B and Figure 5.8B).

5.5. An extensive analysis of co-localization efficiencies between mEGFP-U2AF65 and globin C derivatives

To gain further insight into the mechanisms by which U2AF65 could bind the globin C pre-mRNA, an extensive analysis of co-localization efficiencies between globin C derivatives and mEGFP-U2AF65 under different experimental conditions was carried out. The NE used for these experiments was kept constant so as to be able to compare experiments, it was made from HEK 293T cells that had been transfected with plasmids encoding the mEGFP-U2AF65 and mCherry-U2AF35 fusion proteins. The NE spliced efficiently (Figure 5.10A) and expression of mEGFP-U2AF65 was easily detectable by western blot (Figure 5.10B). The splicing conditions used were altered

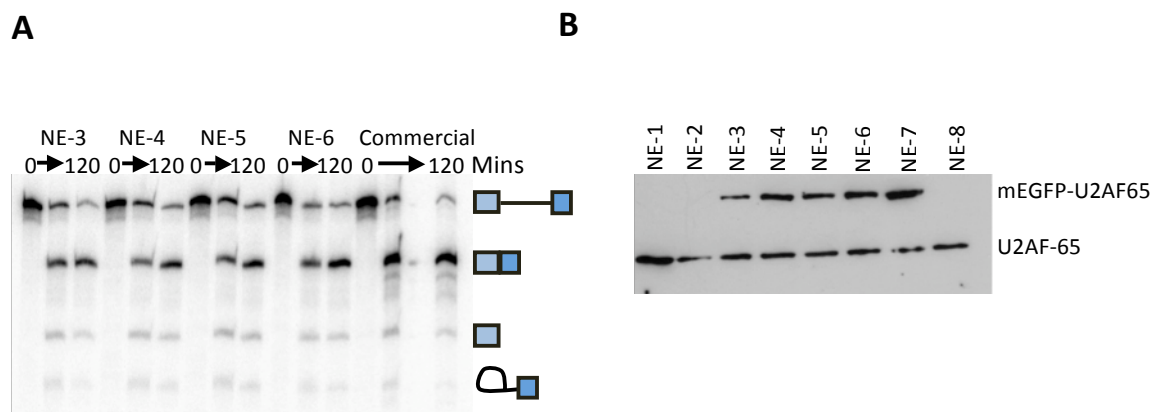


Figure 5.10: Characterization of the NE used for TIRF microscopy experiments. (A) Denaturing urea 6 % polyacrylamide gel of splicing products produced by a series of NE made from HEK 293T cells co-transfected with plasmids encoding mEGFP-U2AF65 and mCherry-U2AF35 (NE-3-6) or commercial NE. **(B)** Western blot of the series of NE made from HEK 293T cells transfected with different amounts of the plasmids encoding mEGFP-U2AF65 and mCherry-U2AF35. The blot was probed using the α U2AF65 antibody (Table 2.3). For the TIRF microscopy experiments, NE-4 was chosen as its expression was high and it spliced efficiently; this NE was made from HEK 293T cells that had been transfected with 2 μ g of each plasmid.

so that they matched those used in standard splicing reactions (Section 2.1.3) and not those previously used for single molecule splicing reactions (Section 2.3.4).

The first transcript assayed was the full-length globin C pre-mRNA. Co-localization efficiency under E complex conditions ranged from 28 % to 21 % and under A complex conditions ranged from 21 % to 18 % (Table 5.3). To assess the effect of U1 snRNP sequestration on co-localization between mEGFP-U2AF65 and globin C, experiments were carried out under E complex and A complex conditions in which 4 μ M α -U1 oligonucleotide was added. Under either E complex (-ATP +U1) or A (+U6 +U1) conditions the addition of the α -U1 oligonucleotide had no effect on co-localization efficiency (Table 5.3). The observed co-localization was specific, as addition of the globin C pre-mRNA to a splicing reaction after both the splicing mixture and RNA had been diluted to the concentration needed for TIRF microscopy did not yield significant levels of co-localization (Table 5.6; Addition post dilution). To try and determine the number of U2AF65 proteins bound to this pre-mRNA the intensity profiles were investigated for the globin C –ATP (a) sample. Unfortunately when the intensity profiles were examined, approximately 62 % of the intensity profiles were not assignable, those that were assignable often had varied fluorescence intensity. Therefore, this data was not analysed in terms of the number of proteins bound. This was unexpected as intensity profiles from previous data obtained using NE made from HeLa cells expressing the mEGFP-U2AF65 fluorescent fusion protein (Figure 5.1) had been easily assignable. Possible reasons for this discrepancy are discussed in the discussion.

Co-localization efficiencies were also assessed for the globin M pre-mRNA. Despite the absence of a 5' SS, co-localization remained at approximately 20 % whether analysed under E or A complex conditions (Table 5.3). This is in agreement with observations indicating spliceosomal related complexes can form on 3' half substrates containing the 3' SS (For example: Michaud and Reed, 1993).

Sample	Total Cy5 spots	Co-localized Cy5 spots	Percentage co-localization
globin C –ATP (a)	348	96	28
globin C –ATP (b)	451	100	22
globin C –ATP (c)	238	48	20
globin C +U6 (a)	674	142	21
globin C +U6(b)	462	84	18
globin C +ATP (c)	429	78	18
globin C –ATP +U1	391	83	21
globin C + U6 +U1	650	129	20
globin M –ATP	515	108	21
globin M + U6	449	103	23

Table 5.3: Efficiency of co-localization between mEGFP-U2AF65 and globin C or globin M.

In contrast to the data for the full length constructs, when a construct lacking any sequence downstream of the BPS, globin C Δ 3' RNA, was used for the same experiments it elicited a different response to the inclusion of the α -U1 oligonucleotide in the reaction. Under E complex conditions, addition of the α -U1 oligonucleotide had no effect on co-localization efficiency (Table 5.4). However, under A complex conditions, addition of the α -U1 oligonucleotide to the reaction led to a drop in co-localization efficiency from 13 % to 5 % (Table 5.4). This was in agreement with the data obtained for the globin C Δ pY, Δ 24-36U/C RNA (Table 5.2), confirming that the U1 snRNP plays a role in recruiting U2AF65 if the 3' SS region is absent. This suggests that the U1 snRNP-dependent recruitment of mEGFP-U2AF65 is not involved in the recognition of the downstream 3' SS.

This set of experiments was also performed on the globin C Δ pY RNA and co-localization again dropped to 5 % under A complex conditions in the presence of the α -U1 oligonucleotide (Table 5.4). Surprisingly under A complex conditions in the absence of the α -U1 oligonucleotide, co-localization efficiency also appeared to drop (Table 5.4). It is unknown why this was observed, but it is likely to be an anomaly, possibly caused by RNA degradation. It is unlikely to be due to the presence of the 3'exon sequence as the globin C Δ pY, Δ 24-36U/C RNA did not show the same loss of co-localization under A complex conditions (Table 5.2).

To assess whether the effect of α -U1 oligonucleotide inclusion was due to the oligonucleotide itself and not due to sequestration of U1 snRNP, the globin M Δ 3' RNA was analysed. This RNA did not have a 5' SS and therefore if the effect of α -U1 oligonucleotide inclusion was due to blocking of the association between U1 snRNP and the 5' SS, it was expected that there would be a loss of co-localization under A complex conditions, in the absence of the α -U1 oligonucleotide. As expected globin M Δ 3' showed 14 % co-localization under E complex conditions, which dropped to 5 % under A complex conditions (Table 5.4). This showed that the drop in mEGFP-U2AF65 co-localization observed in the presence of the α -U1 oligonucleotide is not due to the oligonucleotide itself. Instead the drop in co-localization is due to a lack of U1 snRNP bound to the 5' SS. Interestingly, the U1 snRNP dependent co-localization between mEGFP-U2AF65 and pre-mRNA does not appear to be due to a direct interaction, as no significant co-localization was observed between mEGFP-U2AF65 and the α -U1-Cy5 oligonucleotide (Table 5.6), which binds U1 snRNP (Section 4.4).

Sample	Total Cy5 spots	Co-localized Cy5 spots	Percentage co-localization
globin C -3' –ATP	292	52	18
globin C -3' + U6	356	47	13
globin C -3' –ATP +U1	440	75	17
globin C -3' + U6 +U1	465	25	5
globin M -3' –ATP	381	52	14
globin M -3' + U6	423	21	5
globin C ΔpY –ATP	428	96	22
globin C ΔpY + U6	688	70	10
globin C ΔpY + U6 +U1	415	20	5

Table 5.4: Efficiency of co-localization between mEGFP-U2AF65 and globin derivatives lacking the 3'ss region

Overall these results suggest that there are three ways in which mEGFP-U2AF65 can associate with a pre-mRNA. Firstly there appears to be a “non-specific” interaction under E complex conditions with all RNAs tested. Secondly there appears to be a specific association that is dependent upon U1 snRNP binding the 5' SS and does not require the pY tract. Loss of this interaction in the absence of U1 snRNP binding is only observed in the presence of ATP. This suggests either that the U1 snRNP-dependent recruitment of U2AF65 only occurs in the presence of ATP or that this interaction is lost in the absence of ATP but that the non-specific interaction masks its loss. Thirdly there is a specific interaction that is dependent upon the presence of a pY tract, which is not dependent upon the U1 snRNP.

The “non-specific” association observed between all RNAs and mEGFP-U2AF65 under E complex conditions may have been due to dephosphorylation of U2AF65, leading to non-specific aggregation with RNA. To assess this, co-localization between globin pre-mRNAs and mEGFP-U2AF65 in the presence of Phos-stop phosphatase inhibitor was investigated. Phos-stop was added to the NE, at a concentration of 2X, prior to pre-

incubation of the NE at 30 °C for 30 minutes to deplete ATP. This NE was then added to a splicing reaction lacking ATP and Phosphocreatine and incubated with fluorescent oligonucleotide-labelled pre-mRNA for 30 minutes at 30 °C to allow complex formation. Analysis of co-localization between mEGFP-U2AF65 and either globin C or globin MΔ3', under these conditions, showed that inhibition of phosphatases had no effect on the co-localization efficiency observed (Table 5.5). Hence the “non-specific” co-localization was not caused by dephosphorylation of proteins due to the ATP depletion.

Sample	Total Cy5 spots	Co-localized Cy5 spots	Percentage co-localization
globin C –ATP +PS	511	115	23
globin M -3'–ATP +PS	356	65	18

Table 5.5: Efficiency of co-localization between mEGFP-U2AF65 and globin C or globin M-3' in the presence of Phos-stop phosphatase inhibitor

Sample	Total Cy5 spots	Co-localized Cy5 spots	Percentage co-localization
α-U1-Cy5	332	4	1
Addition post dilution (globin C)	83	2	2

Table 5.6: Efficiency of co-localization in control experiments

5.6. Functional characterization of other fluorescent fusion proteins

The fluorescent fusion proteins used in this study represent only a few of the total cloned by Dr L. P. Eperon. Therefore experiments have been carried out to assess whether some of the other proteins are functional. To assess whether the mEGFP-U5-40K and mEGFP-U2B'' fluorescent fusion proteins were functional, immunoprecipitations from splicing reactions containing radiolabelled globin C pre-

mRNA in the presence of ATP were carried out. Nuclear extracts were produced from HeLa cells expressing the two fluorescent fusion proteins and added to standard splicing reactions containing radiolabelled globin C pre-mRNA. The reactions were then incubated for 30 minutes at 30 °C, to allow spliceosomal complex formation, prior to immunoprecipitation using the α EGFP antibody (Table 2.3). Analysis of the radiolabelled RNAs immunoprecipitated shows that both mEGFP-U5-40K and mEGFP-U2B'' bind pre-mRNA (Figure 5.11A). When the snRNAs immunoprecipitated from NE expressing either mEGFP-U2B'' or mEGFP-U1-A were analysed by northern blot using snRNA riboprobes (Blencowe et al., 1989), it was clear that the two fusion proteins interacted with the expected snRNAs (Figure 5.11B).

Immunoprecipitations of the fluorescent fusion protein, mEGFP-SF2 (SFRS-1) were also carried out. For these immunoprecipitations, standard splicing reactions lacking ATP and Phosphocreatine were set up, using NE prepared from HeLa cells expressing mEGFP-SF2 that had been incubated at 30 °C for 30 minutes to deplete ATP. Under these conditions mEGFP-SF2 immunoprecipitated the globin C pre-mRNA very efficiently (Figure 5.11C). When the immunoprecipitated snRNAs were assessed by northern blot using the snRNA riboprobes it was found that mEGFP-SF2 had immunoprecipitated both U1 snRNA and U2 snRNA (Figure 5.11D). It is unknown whether this is a direct interaction or if it is mediated by pre-mRNA.

As a second way to assess whether mEGFP-SF2 (SFRS-1) and mCherry-SF2 were functional, an *in vivo* assay was designed. SFRS-1 is known to autoregulate its own expression levels through multiple mechanisms (Sun et al., 2010), one of which is through changes in alternative splicing (Lareau et al., 2007; Ni et al., 2007; Sun et al

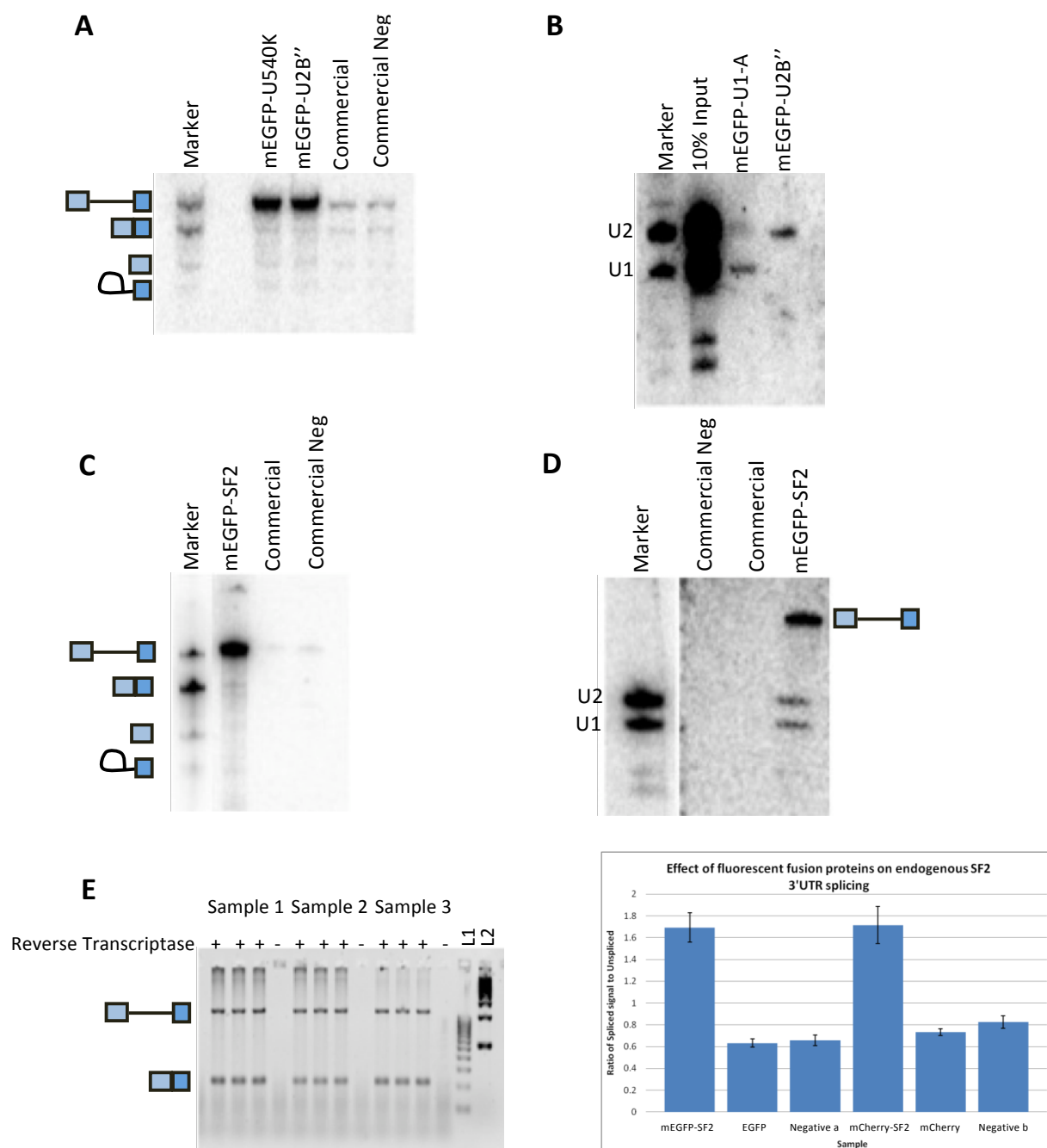


Figure 5.11: Functional assays performed on other spliceosomal factors. (A) Immunoprecipitation of radiolabelled globin C pre-mRNA from standard splicing reactions made using NE containing the indicated fluorescent fusion protein or commercial NE. Immunoprecipitation was performed using the α EGFP antibody (Table 2.3). The sample labelled Neg had no antibody added. (B) Northern blot of the RNAs immunoprecipitated, using the α EGFP antibody, from extract containing mEGFP-U1-A or mEGFP-U2B''. The blot was probed using the riboprobes designed to detect the snRNAs. (C) Immunoprecipitation of radiolabelled globin C pre-mRNA from standard splicing reactions, lacking ATP and phosphocreatine, using the α EGFP antibody. Splicing reactions were made with either NE containing mEGFP-SF2 or commercial NE. Sample labelled Neg had no antibody added. (D) Northern blot performed on the RNAs precipitated by the immunoprecipitation shown in Figure 5.10C. Blot was probed using the riboprobes designed to detect the spliceosomal snRNAs. (E) *In vivo* assay for the function of fluorescent fusion-SF2 (SF2S-1) constructs. The left hand panel shows sample data showing the bands detected for spliced and unspliced intron, L1 refers to the Hyper ladder IV (Bioline), L2 refers to the 1Kb ladder (New England Biolab). The right hand panel shows the results of quantification of the ratio between the spliced and unspliced bands for each set of transfections; error bars represent standard error of the mean for each set of transfections.

2010). SFRS-1 has been shown to activate cryptic introns within the long 3'untranslated region (UTR) of its own pre-mRNA, thus stimulating nonsense-mediated-decay or nuclear retention of its own transcript (Sun et al., 2010). To see whether the mEGFP-SF2 and mCherry-SF2 fluorescent fusion proteins caused an increase in the excision of one of the introns within the 3'UTR, a reverse transcriptase-PCR assay was designed.

For this assay HeLa cells were seeded into 6-well plates and transfected with the mEGFP-SF2, mCherry-SF2, EGFP or mCherry plasmids using Eugene-6. Cells were cultured for ~48 hours before the RNA within the cells was harvested using Tri-reagent (Section 2.2.9). From this RNA, cDNA was then synthesised and analyzed by PCR. The PCR reactions used the oligonucleotides SF2-F and SF2-R (Table 2.1), which had been designed to amplify a fragment of 1136 nt in length, starting at position 2234 nt and terminating at position 3370 nt, relative to the transcription start site. These sites had been chosen as they spanned the cryptic intron at position 2305 nt to 3225 nt (Sun et al., 2010). Amplification using these primers yielded two products; the first was from the un-spliced transcript and was 1136nt long and the second was from the spliced transcript and was 216nt long. The PCR reactions were performed using Red-Taq polymerase and were run on a 1.5 % agarose gel, prior to staining with ethidium bromide and imaging. Representative data is shown in figure 5.11E. To determine whether the mEGFP-SF2 and mCherry-SF2 fluorescent fusion proteins affected the cryptic intron splicing efficiency, the ratio between spliced and unspliced product was calculated. Each fluorescent protein was transfected into HeLa cells at least three times, and each cDNA sample made from the transfections was then PCR amplified three times to control for PCR variability. An average ratio was then calculated for

each transfection from the three PCR reactions, and this average from each transfection was then used to calculate an average for each fluorescent protein. As can be seen in Figure 5.11E and Table 5.7, transfection of plasmids encoding mEGFP-SF2 or mCherry-SF2 increased the amount of spliced product relative to the unspliced product dramatically, whilst the EGFP and mCherry controls had no effect, showing that the fluorescent fusion protein versions of SF2/ASF behaved as expected in splicing regulation.

Sample	Level of Spliced Product	Standard Deviation	Standard Error
mEGFP-SF2	1.692507385	0.329870789	0.134669
EGFP	0.633028978	0.073472597	0.036736
Negative a	0.656551746	0.109990329	0.049189
mCherry-SF2	1.715939379	0.296926212	0.17143
mCherry	0.73302696	0.056612297	0.032685
Negative b	0.825824662	0.097454326	0.056265

Table 5.7: The effect SF2-fusion proteins on SF2 3'UTR splicing. Sample names refer to the source of the cDNA; cDNA was synthesized from cells transfected with plasmids encoding the indicated fusion proteins. "Level of spliced product" refers to the intensity of the PCR product from the spliced form, when the un-spliced form is normalised to one.

Summary

Through this work the interactions between U2AF65 and derivatives of the globin C pre-mRNA have been investigated. Instead of the expected, simple, model in which U2AF65 can only bind the pY tract of the intron, it appears that a pre-mRNA with two exons and one intron is bound by U2AF65 using three mechanisms. The first is a "non-specific" mechanism in which U2AF65 binds the pre-mRNA in the absence of ATP, the second is a U1 snRNP-dependent interaction with the pre-mRNA that is not dependent on the pY tract and the third is the expected interaction between U2AF65 and the pY tract.

Chapter 6. Discussion

6.1 The effect of DNase on globin C pre-mRNA splicing

6.2 The role of U1 snRNP in 5' SS selection

6.3 Interpretation of data collected from TIRF microscope

6.4 Possible explanations for the intensity profiles of mEGFP-U2AF65

6.5 Mechanisms of association of mEGFP-U2AF65 with globin C pre-mRNA

6.6 Concluding remarks

The purpose of this research was to analyse the mechanism by which the factors that first recognise and bind the splice sites, the U1 snRNP and U2AF, mediate splice site selection. Is their binding the key to subsequent use of a sequence as a splice site? Single molecule methods have much to offer as tools for such work but the novelty of the methods means that substantial preliminary characterization and validation are required. Some of the observations made in this phase are of interest, and one of these, the effect of DNase, needs some discussion.

6.1. The effect of DNase on globin C pre-mRNA splicing

The observation that pre-incubation of NE with DNase inhibited splicing, was unexpected. The data shows that when NE is pre-incubated with active DNase it can digest large pieces of DNA that are detected as a single band on a 10 cm long denaturing urea 6 % polyacrylamide gel (Figures 3.15A and 3.15B). The DNase does not appear to be contaminated with RNase, as none of the snRNAs within the NE are affected by the pre-incubation with DNase, suggesting that digestion of the large DNA molecule is responsible for the observed inhibition of splicing. This is remarkable as there has never been any previous suggestion that DNA is involved in splicing.

There are two possible explanations for the effect of DNase on splicing. Firstly the DNA fragments digested by DNase may play a direct role in spliceosome assembly or structure. However, this seems unlikely as DNA has never been reported to be part of the purified spliceosome (For example: Jurica et al., 2002; Hartmuth et al., 2002; Makarov et al., 2002; Jurica et al., 2004; Deckert et al., 2006; Behzadina et al., 2007). The observation that the

DNA has a smeary appearance when run on a 20 cm long denaturing urea 6 % polyacrylamide gel (Figure 3.12), suggests that it is heterogeneous. It seems unlikely that a collection of DNA molecules of different lengths and sequences would play a role in splicing. The second possibility is that the DNA may sequester splicing regulatory factors, possibly indicating, that DNA is required for an active splicing NE. It is possible that the heterogeneous nature of the DNA fragment is as a result of the nuclear extraction procedure; the splitting of the cells may lead to the genomic DNA within the cells being sheared. This sheared DNA may be naked or wrapped up in chromatin and may be able to bind spliceosomal factors such as hnRNP A1, Srp-20, SF2/ASF and PSF, all of which have been shown to bind DNA or chromatin (Fiset et al., 2001; Donev et al., 2002; Loomis et al., 2009; Akhmedov and Lopez, 2000).

To investigate this phenomenon further, it would be important to determine whether pre-treatment of NE with DNase has an effect on the splicing efficiency of other pre-mRNA transcripts. This would allow us to understand whether the effect is a general effect on all splicing, or if it is specific to a subset of transcripts. Further to this it would be interesting to investigate the point at which pre-treatment with DNase effects splicing. This could be done by investigating the effect of DNase on the spliceosomal assembly pathway, using native LMP agarose gel electrophoresis. This may give an indication as to the point at which the treatment is having an effect and the factors involved. To determine whether the DNA in NE is single or double stranded, NE could be incubated with single strand-specific nucleases, such as Exonuclease III or S1 Nuclease. This may also give an indication as to the identity of the factors involved.

The identification of exactly which factors are affected by DNase treatment may prove difficult. This is because if the fragment is a heterogeneous mix of randomly sheared genomic DNA, an affinity purification approach, with a selected sequence, would only show you a subset of the total proteins bound to the DNA. However, it may be possible to purify the DNA band from a gel and ligate onto each DNA fragment an adapter sequence that is labelled with a biotin group. This pool of biotinylated DNAs, with unknown sequences, could then be used to affinity purify proteins that bind the DNA in NE, possibly leading to the identification of some candidate proteins to explain the effect of DNase on splicing.

6.2. The role of U1 snRNP in the position effect on 5' SS selection

The U1 snRNP has long been thought to be able to associate with the 5' SS of a pre-mRNA (Lerner et al., 1980; Rogers and Wall, 1980). Subsequently it was found to be needed for splicing (Padgett et al., 1983; Kramer et al., 1984) and to bind the 5' SS sequence (Mount et al., 1983). However, in pre-mRNA containing more than one 5' SS, the role of U1 snRNP has been far harder to interpret. Although the strength of base-pairing between the U1 snRNA and 5' SS is clearly important, the correlation between predicted strength of base-pairing and actual use breaks down when base-pairing is predicted to be moderate. The discovery of protein factors that can modulate U1 snRNP association and compete with it gives some indication of how the affinity of U1 snRNP for the 5' SS can be altered by other factors (Section 1.3). The observations indicating that sequence can influence the affinity of U1 snRNP for 5' splice sites led to the formation of hypotheses based on affinity,

whereby 5' SS selection occurred via a single U1 snRNP selecting which 5' SS to use.

Models based upon affinity to explain 5' SS selection included a model in which U1 snRNP was sequestered at another point on the pre-mRNA and from that point selected the 5' SS to be used, and a model based upon kinetics whereby whichever 5' SS was bound first would be used (Cunningham et al., 1990; Eperon et al., 1986; Figure 1.3; Section 1.4).

However, the models based solely upon the affinity of U1 snRNP for the 5' SS could not explain the observed position effect, i.e., that when two 5' SS of comparable predicted strength are competing for use in splicing, the downstream 5' SS is heavily favoured (Reed and Maniatis, 1986; Eperon et al., 1993; Hicks et al., 2010; Yu et al., 2008). If selection is based upon affinity the preference for the downstream 5' SS is very difficult to explain. An alternative hypothesis was proposed, based upon ribonuclease protection assays, in which the 5' splice sites can be occupied by two U1 snRNPs independently. This means that the spliceosome has to choose between the multiple U1 snRNPs bound to a transcript, but the mechanism leading to the use of the downstream 5' SS was unknown (Eperon et al., 1993; Eperon et al., 2000).

Two U1 snRNPs can bind a pre-mRNA containing two 5' SS

The finding that two U1 snRNPs can be directly observed bound to one pre-mRNA under E complex conditions clearly shows that a model based only on affinity of U1 snRNP for a 5' SS can be rejected. In contrast the model whereby the two 5' SS on the pre-mRNA can be occupied independently by two U1 snRNPs can be accepted. In this model the probability of occupation of each of the two 5' SS determines which one is selected to be used in

splicing. As part of this model the upstream 5' SS would only be selected if the downstream 5' SS is unoccupied, whilst the downstream 5' SS would be selected whether or not the upstream 5' SS is occupied (Figure 1.4). In cases where the probability of occupation is low, selection is likely to occur based on affinity, whereby whichever of the two 5' SS is occupied will be used in splicing.

This model, is backed up by a number of observations: (a) ribonuclease protection assays have indicated that two 5' SS on a single pre-mRNA can be protected in a U1 snRNP dependent manner (Nelson and Green, 1988; Eperon et al., 1993); (b) when U1 snRNP has been cross-linked to the 5' SS using psoralen, the cross-linking efficiency for the downstream 5' SS correlates with its use whilst the cross-linking efficiency of the upstream 5' SS does not, which fits the model in that if the downstream 5' SS is bound by U1 snRNP then it is used, whilst the upstream is not always used when bound (Eperon et al., 2000); (c) in competition assays, significant use of the upstream 5' SS is only observed when the downstream 5' SS is considerably weaker than the upstream, meaning that there would be a higher chance that at some points only the upstream 5' SS is occupied (Eperon et al., 1993); (d) high concentrations of SFRS-1 stimulate binding of U1 snRNP to both 5' SS which, according to the model, would stimulate the use of the downstream 5' SS exactly as is observed (Eperon et al., 1993; Eperon et al., 2000). (e) pre-incubation of NE with a 2'O methyl oligonucleotide designed to sequester the U1 snRNP can shift 5' SS selection, from almost exclusively downstream to partially upstream (Figure 4.9), As predicted if lower concentrations of free U1 snRNP reduce the level of double occupancy.

This last point raises the interesting possibility that the concentration of U1 snRNP in a nucleus may play a role in 5' SS selection. Interestingly, experiments investigating an oligonucleotide that was designed to tether the U1 snRNP to the final exon of a gene have shown that the oligonucleotide actually sequesters U1 snRNP, and alters the pattern of alternative splicing so that an exon can be excluded from the final mRNA (Vickers et al., 2011). This suggests, in agreement with the data presented in Figure 4.9, that the concentration of free U1 snRNP in the NE is of critical importance for determining the choice of 5' SS. The idea that levels of free U1 snRNP may influence 5' SS selection is backed up by three points.

First, when the levels of non-coding-RNAs were investigated in eleven human tissues the levels of U1 snRNA was found to vary (Castle et al., 2010), which may lead to alterations in the pattern of alternative splicing.

Second, studies in which the level of U1 snRNP protein factors have been reduced by siRNA mediated knockdown, have shown an effect on alternative splicing patterns.

Knockdown of U1-70K or U1-C in *D.melanogaster* led to a shift in 5' SS selection, with the upstream 5' SS being selected on the construct tested (Park et al., 2004); this may have been due to reduced levels of functional U1 snRNP leading to lower occupancy of the 5' splice sites. Knockdown of U1-C in *D. rerio* also led to alterations in the pattern of 5' SS selection, in which 5' SS selection was shifted both upstream and downstream according to the pre-mRNA (Rosel et al., 2011). This variation in effect suggests multiple factors are influencing 5' SS selection. Interestingly poly-uridine tracts were associated with those

pre-mRNAs influenced by U1-C knockdown, suggesting that other factors such as Tia-1 may be influencing the outcome. This is likely as Tia-1 is known to enhance the binding of U1 snRNP to 5' SS through an interaction with U1-C (Forch et al., 2002).

Third siRNA-mediated depletion of the core Sm proteins B/B' also led to the conclusion that levels of U1 snRNP were important in 5' SS selection (Saltzman et al., 2011). The Sm proteins B/B' are part of the Sm protein complex that is assembled onto the Sm site of the U1, U2, U4 and U5 snRNAs in an SMN dependent manner (Neuenkirchen et al., 2008). The inclusion of a cassette exon in the B/B' RNA containing a premature stop codon was observed to require high levels of B/B', this exon had a weak 5' SS and when levels of B/B' were high its recognition by U1 snRNP was increased. This was hypothesised to be because, when the levels of B/B' are high, the levels of functional U1 snRNP are high and as such the probability of recognition of the weak 5' SS in the B/B' cassette exon is increased (Saltzman et al., 2011)

Interestingly the alterations in the expression levels of the U1 snRNA gene may lead to greater use of 5' SS defined by U1 snRNPs made from lesser U1 snRNAs. Expression of non-canonical U1 snRNAs has been detected at low levels. These U1 snRNAs can have different 5' termini and could recognise different 5' SS sequences (Lund and Dahlberg, 1987; Kyriakopoulou et al., 2006). Although expressed at low levels, some of these U1 snRNAs have been found to associate with Sm proteins in RNA-protein complexes (Kyriakopoulou et al., 2006). This suggests they may be able to form functional U1 snRNPs that can select for non-canonical 5' SS sequences. Lower levels of the normal U1 snRNA

would mean that there would be a higher probability of a non-canonical U1 snRNA-containing U1 snRNP being used.

In cases where more than one 5' SS is competing for use, selection of the upstream 5' SS ahead of the downstream 5' SS can be achieved in a number of ways in addition to the alterations in the level of functional U1 snRNP. There are examples whereby proteins such as hnRNP A1 (Eperon et al., 2000) and hnRNP H (Buratti et al., 2004) have been found to be able to compete with U1 snRNP for binding of the 5' SS. If an affected 5' SS was in competition with another it would not be used as there would be no U1 snRNP present. Other experiments have shown that hnRNP A1 does not always interfere with U1 snRNP binding and it has been suggested that hnRNP A1 can also loop out an alternatively spliced exon through homodimerisation of two or more hnRNP A1 proteins bound to two high affinity sites either side of the exon (Blanchette and Chabot, 1999; Nasim et al., 2002). This work was carried out using an alternative exon as a reporter but it is also possible that hnRNP A1 is able to loop out 5' SS and as such select against them. PTB has also been suggested to loop out alternative exons (Oberstrass et al 2005) and so may also be able to alter 5' SS selection by the same mechanism.

RNA structure around a 5' SS has also been shown to play a role in determining whether it is selected or not. Sequestration of a 5' SS within a stem-loop structure both in native pre-mRNAs (Blanchette and Chabot, 1997; Abbink and Berkhout, 2008; Singh et al., 1997; Loeb et al., 2002; Jiang et al., 2000) and when artificially added (Solnick and Lee, 1987; Eperon et al., 1988) have been shown to effect 5' SS selection. This sequestration of the 5' SS

leads to a loss of U1 snRNP binding (Blanchette and Chabot, 1997; Jiang et al., 2000; Singh et al., 2007).

In some cases selection of an upstream 5' SS may also occur after U1 snRNP association with the 5' SS. The proteins HMGA1a, PSI and PTB are all known to interact with U1 snRNP and interfere with its recognition and use in splicing (Ohe and Mayeda, 2010; Labourier et al., 2001; Sharma et al., 2011).

One U1 snRNP is removed from pre-mRNA containing two 5' SS during the E complex to A transition

E complex is the first specific spliceosomal complex to form on a pre-mRNA (Jamison et al., 1992; Michaud and Reed, 1991; Michaud and Reed, 1993) and is required for A complex formation (Jamison et al., 1992). Analysis of the relative positions of the 5' and 3' SS within E complex showed that they are in close proximity to one another, suggesting that they must be connected at this early stage in splicing (Kent et al., 2002; Staknis and Reed, 1994b). This agrees with the finding that E complex formation is enhanced by the presence of both the 5' and 3' SS sequences within a pre-mRNA (Michaud and Reed, 1993) and the finding that U1 and U2 snRNPs are in close proximity in E complex (Donmez et al., 2007).

With this in mind, it is interesting that in cases where there are multiple 5' or 3' SS, a commitment as to which of the available splice sites a spliceosome will use is not made until A complex (Lim and Hertel, 2004; Kotlajich et al., 2009).

Even though these results seem anomalous, they are borne out by our observation that a U1 snRNP is lost from the C174C and Ad1 CC pre-mRNAs, which contain two 5' SS, between the E complex and A complexes. Reconciling these results does not mean that both U1 snRNPs and 5' SS make contacts with the 3' SS. It is possible that, even in E complex, the downstream 5' SS is already in contact with the 3' SS because of an even earlier U1 snRNP-induced change in the structure of the spliceosome and that in the assays of Kotlajich et al. (2009) only a small proportion of pre-mRNA molecules have been bound by U1 snRNP. This interpretation is consistent with our model, as the upstream 5' SS is a cryptic 5' SS and the downstream 5' SS is dependent upon the presence of Tra2 β .

The mechanistic reason for the loss of one of the U1 snRNPs is unknown; however, there are some obvious possibilities. Although splicing itself probably enhances the export of mRNA from the nucleus (Valencia et al., 2008), the recognition of a 5' SS in a pre-mRNA that is not spliced is known to inhibit export (Chang and Sharp, 1989; Legrain and Rosbash, 1989; Huang and Carmichael, 1996). Deletion of the 5' SS, or use of mutant strains deficient in U1 snRNP, suggested that this inhibition was due to the binding of U1 snRNP to the 5' SS (Legrain and Rosbash, 1989). Therefore, it is likely that the presence of a U1 snRNP on a transcript inhibits export of the mRNA from the nucleus. In cases where there are two 5' SS splice sites in a pre-mRNA, if the unused U1 snRNP were not removed from the pre-mRNA during splicing it would cause the inhibition of export from the nucleus.

Removal of the unused U1 snRNP may also be necessary to prevent interference with spliceosome formation and regulation at a later point. It may also prevent subsequent

rounds of splicing occurring on the same pre-mRNA. If a U1 snRNP is left associated with the pre-mRNA after splicing, it might stimulate the formation of a second spliceosome which could splice to a previously unrecognised 3' SS, creating a novel mRNA.

Interestingly, in some cases more than one round of splicing (intrasplicing) have been observed to be required for correct mRNA formation (Parra et al., 2008).

The mechanism by which the unused U1 snRNP is removed from the pre-mRNA is another interesting question. The dependence on the addition of ATP and subsequent A complex formation suggests that removal of the unused U1 snRNP may be dependent on an ATP-dependent enzyme, such as an ATP-dependent RNA helicase. Helicases are heavily involved in the regulation and proofreading of spliceosome assembly (Bleichert and Baserga, 2007; Smith et al., 2008; Wahl et al., 2009). A role for a helicase is appropriate because the interaction between the U1 snRNP and 5' SS is partially dependent on the 5' terminus of U1 snRNA base-pairing with the 5' SS, and it is likely that for removal of the unused U1 snRNP this base-pairing would need to be disrupted. One candidate for the removal of the unused U1 snRNP is the RNA helicase P68, which has been implicated in the unwinding of the U1 snRNP-5' SS interaction and has been shown to be necessary for the pre-spliceosome to spliceosome transition (Liu, 2002). It is likely to be associated with the spliceosome at the correct point and might be able to unwind the base-pairing between U1 snRNA and an unused 5' SS.

Although a yeast protein, prp28 has been shown to be needed for the disruption of the interaction between the U1 snRNP and 5' SS (Staley and Guthrie, 1999; Chen et al., 2001),

it is unlikely to influence removal of the unused U1 snRNP. This is because prp28 is part of the U5 snRNP (Stevens et al., 2001) and its human homologue, U5-100K, is associated with both the U5 snRNP and the U4/U6·U5 tri-snRNP (Teigelkamp et al., 1997). Therefore it is not associated with the spliceosome until after the point at which the unused U1 snRNP is removed.

As the point in spliceosomal assembly at which the unused U1 snRNP is removed from the pre-mRNA is between E and A complex, and its dependency on a downstream 3' SS, it is also possible that U1 snRNP removal may be linked with some of the contacts that are formed across the intron at this point. The interactions of yeast Prp5 with the U1 snRNP and U2 snRNP are required for A complex formation (Xu et al., 2004). Although speculative, it is possible that proteins such as prp5 are able to signal to other unknown factors to remove the unused U1 snRNP.

The effect of two 5' SS very close together on the number of U1 snRNPs

Surprisingly, the pre-mRNA C24C, which contains two 5' SS very close together and is able to splice to either 5' SS at a low level (Figure 4.3A), did not show the loss of one U1 snRNP. Instead a loss of co-localization was observed (Table 4.4). This could be the result of two possibilities. First, it could be due to degradation of the C24C pre-mRNA, whilst it was being incubated as part of a splicing reaction for analysis by TIRF microscopy. Although possible, this explanation seems unlikely because the +AA and +U6 samples showed low levels of co-localization whilst none of the other samples on other pre-mRNAs appeared to suffer from degradation. A second possibility is that both U1 snRNPs associated with

the C24C pre-mRNA are removed between E and A complexes. This may be due to the close proximity of the two 5' SS, making it difficult for the splicing machinery to make a distinction between them and remove only one. It is possible that the presence of the two U1 snRNPs under E complex conditions allows progression to A complex but that at some point both are removed, leading to the observed inhibition of the A to B complex transition (Figure 4.10A). This loss of both U1 snRNPs does not inhibit A complex formation, but may lead to an inability to define the 5' SS in B complex, leading to the tri-snRNP not being added. One interesting possibility is that the phosphorylation of Prp28, which is required for tri-snRNP addition to the spliceosome (Mathew et al., 2008) may require the presence of U1 snRNP bound to a 5' SS.

Previously, it has been shown that RNaseH mediated cleavage of the 5' terminus of U1 snRNA caused an increase in splicing efficiency of C24C (Eperon et al., 1993). This interesting result can be explained by hypothesising that U1 snRNPs lacking their 5' terminus can bind pre-mRNA, but due to the resulting decrease in affinity for the 5' SS, they are unlikely to occupy both 5' SS at once. Therefore the pre-mRNA would be likely to only have one U1 snRNP associated at any time and as such would no longer be inhibited. Interestingly this indicates that dual occupancy by U1 snRNPs may be inhibitory to splicing when the two 5' SS are below a minimum distance.

A new model for 5' SS selection

The mechanism by which the downstream 5' SS is selected instead of the upstream 5' SS is still unknown. Although, as discussed, helicases are likely to be involved in the removal of

the unused U1 snRNP, it is still not understood how and why the downstream 5' SS is selected. It is possible that regulatory elements within the region between the two 5' SS influence which of the two 5' SS are chosen, but the efficient splicing of both M174C and C174M suggests that they are not the only factors that lead to selection of the downstream 5' SS.

Recently it has been suggested, in relation to silencer function, that the downstream 5' SS may be selected because it is closer, i.e. the downstream 5' SS has a higher probability of coming into contact with the 3' SS than the upstream 5' SS by three-dimensional diffusion (Yu et al., 2008). It is known that the position effect on 5' SS selection is dependent on the number of nucleotides separating the two 5' SS. Two strong 5' SS separated by more than 45nt show a strong preference for the downstream 5' SS, both *in vivo* and *in vitro* (Cunningham et al., 1991). To test whether this difference was due to differences in the probability that either of the two 5' SS will come into contact with the 3' SS by three-dimensional diffusion, Dr A. Hudson modelled 5' SS selection when two 5' SS are separated by different numbers of nucleotides using a method based upon a freely-jointed chain at thermodynamic equilibrium (Dr A Hudson, personal communication; Rippe, 2001). An increase in the separation of the two 5' SS from 25nt to over 200nt led to a very modest predicted increase in downstream 5' SS use of ~54 % to ~64 % (Dr A Hudson, personal communication; Appendix 1). This clearly does not agree with the observations made from experimental data, in which separation of the two 5' SS by more than 45nt leads to almost 100 % use of the downstream 5' SS (Cunningham et al., 1991; Appendix 1). Although the effects of spliceosomal factors and RNA structure were not

taken into account for this modelling, the results suggest that 5' SS selection cannot be based solely on there being different probabilities of interaction between the 5' SS and 3' SS.

Subsequent to this discovery, a new model has been developed to explain the observed preference for a downstream 5' SS (I.C. Eperon. personal communication). In this model the U1 snRNP bound to the downstream 5' SS stimulates the binding of proteins to the pre-mRNA upstream of it and alters the physical behaviour of the chain. Purified U1 snRNP is known to be able to enhance the binding of SR proteins such as SF2/ASF (Jamison et al., 1995) and in some cases to recruit co-activators such as SRM-160 (Blencowe et al., 1998), which can interact with SR proteins (Longman et al., 2000). It was proposed that the U1 snRNP bound to the downstream 5' SS stimulated the binding of and propagation of SR proteins and co-activators to the pre-mRNA in a polar manner. This assembly of proteins may well be enhanced by the presence of enhancer sequences close to the downstream 5' SS and inbetween the two 5'splice sites. This hypothesis fits well with the finding that binding of U1 snRNP enhances the binding of U2AF65 to the upstream 3' SS, as this enhancement may be mediated by SR proteins (Hoffman and Grabowski, 1992; Wu and Maniatis, 1993; Staknis and Reed, 1994b). Interestingly protein binding in a polar manner is not unheard of; binding of the splicing factor hnRNP A1 in a polar manner has been observed (Okunola and Krainer, 2009).

This binding of SR proteins and coactivators may facilitate exon definition and block use of the upstream 5' SS. One way in which binding of proteins may select for the downstream

5' SS is through sterically inhibiting the U1 snRNP bound to the upstream 5' SS. However, another possibility is that binding of these proteins may make the pre-mRNA more rigid. Modelling of the effect that rigidification of the pre-mRNA between the two 5' SS would have on 5' SS selection showed that the downstream 5' SS would be heavily favoured (Dr A Hudson, personal communication; Appendix 1). This model shows a better fit with the experimentally obtained data.

In summary we hypothesise, that binding of the U1 snRNP to the downstream 5' SS may cause increased rigidity of the exonic regions upstream of the 5' SS and prevent competition from the upstream 5' SS. Interestingly this hypothesis suggests a new mechanism by which trans-acting factors such as hnRNP A1 may act to influence 5' SS selection. They may act to introduce points of flexibility within the pre-mRNA, stimulating the use of the upstream 5' SS.

At the moment the hypothesis of increased exon rigidity by propagation of protein binding remains untested. However, it could be investigated in a number of ways; firstly the effect of replacing sections of the RNA with nucleotide-free linkers to inhibit the propagation of protein binding could be investigated. It is known that addition of nucleotide-free linkers into an intron does not inhibit splicing (Pasman and Garcia-Blanco, 1996), but would it alter the selection of 5' SS? It may also be possible to investigate the effect of introducing secondary structure into the region between the two 5' SS in an attempt to introduce bends into the pre-mRNA and possibly stimulate use of the upstream 5' SS. The effect of U1 snRNP and the presence of the 5' SS on the proteins

bound to the region upstream of the downstream 5' SS could also be investigated using affinity precipitation of the pre-mRNA followed by characterization of the proteins bound, which may allow assessment of whether SR protein and co-activator protein binding has been stimulated. Region-selective labelling with modified nucleotides followed by ultra-violet cross-linking would also allow us to determine the effect of a 5' SS on the proteins bound upstream of the 5' SS.

The increased rigidity model may also allow insight into the mechanism by which spliceosomal factors that inhibit 5' SS use after U1 snRNP association act. Both HMGA1a and PSI are able to inhibit use of a 5' SS after binding of the 5' SS by U1 snRNP (Ohe and Mayeda, 2010; Labourier et al., 2001). Both proteins interact with the U1 snRNP through the U1-70K protein. The U1-70K protein is known to be able to interact with SR proteins (Wu and Maniatis, 1993; Kohtz et al., 1994; Cho et al., 2011). Therefore it is possible that by binding U1-70K, HMGA1a and PSI are able to inhibit the propagation of SR proteins that is normally caused by binding of U1 snRNP to the 5' SS. To investigate whether this is the case a HMGA1a binding site could be introduced close to the downstream 5' SS of C174C and its effect on 5' SS selection and the binding of SR proteins could be assessed.

Overall the findings obtained from the analysis of U1 snRNP binding to pre-mRNA constructs containing multiple 5' SS by TIRF-microscopy and modelling allows the proposal of a new model for 5' SS selection. In this, if two 5' SS are occupied by U1 snRNP use of the downstream 5' SS is enforced by a increased rigidity of the exon and subsequent to

selection the unused U1 snRNP is removed from the pre-mRNA by an unknown mechanism (Figure 6.1).

6.3. Interpretation of data collected from TIRF microscope

In order to draw the conclusions made in section 6.2 the data obtained from the analysis of the TIRF-microscopy experiments had to be interpreted.

About 10 % of spots had two-step intensity profiles either in the absence of pre-mRNA or in the presence of a single 5' SS. Their existence was attributed to weak homodimerization of U1-A (Klein Gunnewick et al., 2000; Varani et al., 2000). These were not factored into the analysis of the data for pre-mRNAs containing two 5' SS. One reason for this is that the proportion of U1 snRNPs bound to a pre-mRNA with two strong 5' SS show a three and four step bleaching profile would be very low. Although in some cases such three or four-step intensity profiles were observed, their numbers were too low to draw any conclusions and they were not incorporated into the analysis.

It is also well documented that when carrying out this type of analysis the proportion of spots with higher numbers of bleaching steps will be underestimated (for example Ulbrich and Isacoff, 2007). As the number of fluorophores present within a spot increase, it becomes more difficult to determine the number of bleaching steps; steps can be lost in the background (Ulbrich and Isacoff, 2007). Loss of the fluorophore signal in the background can occur for a number of reasons. Fluorophore bleaching is essentially a random process and therefore there is a possibility that two fluorophores will bleach within one or two time frames. This means that resolution of these two bleaching steps

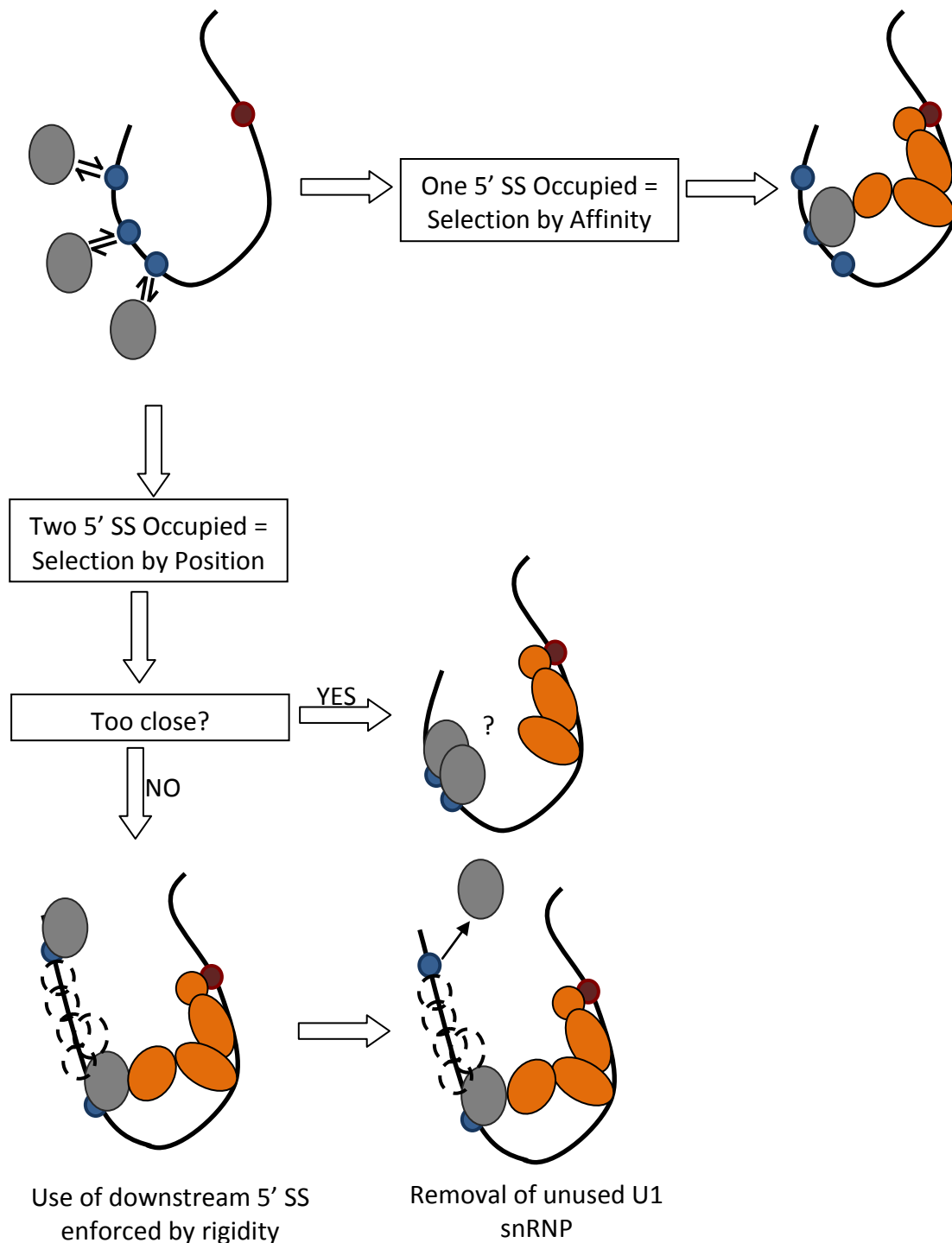


Figure 6.1: A model for the selection of 5' SS. A pre-mRNA with potential 5' SS (blue circles) is bound by U1 snRNP (grey) independently; if only one 5' SS is occupied by U1 snRNP then it is spliced to the 3' SS (red circle), if both 5' splice sites are occupied, one must be selected prior to splicing. Components bound to the 3' SS are shown in orange.

may be limited by the time resolution available (Ulbrich and Isacoff, 2007). Fluorophores with low intensity can also be lost in the background of the sample. The intensity of a fluorophore can vary because of variations in the quantum yield of individual fluorophores. It can also vary due to the distance a fluorophore is from the surface of the slide according to the intensity of the evanescent field. The orientation of a fluorophore relative to the laser may also alter the intensity of the fluorophore; if a fluorophore is 'stuck' in an orientation that does not align with the orientation of the polarised field, then the efficiency of excitation will be low.

All of these factors may lead to fluorophores not being detected, causing an underestimation of the proportion of two-step intensity profiles. We have also assumed that the ratio obtained from the fluorescent western blot gave an accurate measurement of the relative levels of fluorescent and non-fluorescent U1-A that are part of the U1 snRNP in the NE. It was assumed that mEGFP-U1-A had the same RNA binding properties as U1-A, due to the observation that mEGFP-U1-A can associate with the U1 snRNA. However, this is unknown and may affect the proportion of U1 snRNPs containing mEGFP-U1-A and endogenous U1-A. We have also assumed that all of the mEGFP-U1-A is fluorescent; fluorophores may not fluoresce for a number of reasons. The protein may be mis-folded, or folded into an inactive state. Finally we do not know the proportion of fluorophores that are bleached before the data is collected. Previous experiments have estimated the latter to account for up to approximately 10 % of all fluorophores (Cherny et al., 2010).

In conclusion it is not certain that the ratio of fluorescent to non-fluorescent U1-A are correct, and it is unlikely that all of the 'fluorescent' protein has been detected.

Nonetheless, the combination of the results seen with various substrates expected to bind one or two U1 snRNPs argues strongly that the interpretations given in section 6.2 are sound.

6.4. Possible explanations for the intensity profiles of mEGFP-U2AF65

As described in section 5.5, experiments carried out in which co-localization between mEGFP-U2AF65 and derivatives of the globin C pre-mRNA was assessed gave uninterpretable intensity profiles. This unfortunately meant we were unable to determine exactly how many U2AF65 were bound to the globin C pre-mRNA, although from the mutational analysis it appears to be more than one. Possible reasons for this inability to determine bleaching events in the intensity profiles of the mEGFP-U2AF65 spots are given below.

Firstly, it is possible that part of the reason is that the NE used was made from cells co-transfected with mEGFP-U2AF65 and mCherry-U2AF35. A previous report has observed Fluorescence Resonance Energy Transfer (FRET) between an EYFP-U2AF65 and ECFP-U2AF35 (Chusainow et al., 2005). Since FRET can occur between the mEGFP and mCherry fluorophores (Albertazzi et al., 2009), it is likely that FRET is occurring between the mEGFP-U2AF65 and mCherry-U2AF35 fusion proteins. This would lead to fluctuations in the intensity of the mEGFP-U2AF65 fluorophores, as some of them would be transferring energy to the mCherry fluorophore, instead of releasing it as a photon. This hypothesis is

supported by the observation that the mEGFP-U2AF65 and mCherry-U2AF35 fluorophores co-localise, at the single molecule level (Figure 5.5D).

Another possible factor contributing to the observed intensity profiles is the cell type from which the NE was made. The preliminary experiments (Section 5.1) were carried out using NE made from HeLa cells, whilst the later experiments (Section 5.5) were carried out using NE made from HEK-293T cells. It is possible that there are differences in the post-translational modification of mEGFP in the two cell lines and that these may alter the fluorescence properties of the fluorophore.

It is also possible that the level of Argon in the Argon-Ion 488 nm laser was low; this would lead to the intensity of the light being produced by the laser being more variable than it should be. As such the intensity profile of the mEGFP-U2AF65 fluorophore would be difficult to interpret.

6.5. Mechanisms of association of mEGFP-U2AF65 with globin C pre-mRNA

As described in Section 5.5, it appeared that mEGFP-U2AF65 associated with globin C and variants thereof, in three distinct ways. These three mechanisms are summarised in Figure 6.2.

The first mechanism appeared to be “non-specific”, because it was not dependent on the pY tract of globin C or the association of U1 snRNP with the 5' SS. Co-localisation was observed even in the presence of the α -U1 oligonucleotide, which prevents E complex formation (Figure 3.4C), suggesting that mEGFP-U2AF65 bound globin C in H complex.

Previous studies have detected U2AF65 in H complex purified by affinity chromatography (Sharma et al., 2005). This association too may have been non-specific. The non-specific association may occur naturally, or it may be caused by the over-expression of mEGFP-U2AF65, which may in turn lead to the protein associating with the pre-mRNA, essentially at random. However, when the ratio between mEGFP-U2AF65 and endogenous U2AF65 was calculated based on ultra-violet cross-linking (Figure 5.5B), it was observed that the ratio was 1 mEGFP-U2AF65: 2.5 endogenous U2AF65, making it unlikely that overexpression is the cause of the co-localization.

The second mechanism is dependent upon the association of U1 snRNP with the 5' SS. This association was only detected under A complex conditions, which contain ATP and Phosphocreatine. This may either indicate that the U1 snRNP dependent association of mEGFP-U2AF65 is dependent upon ATP, or that the presence of ATP in the reaction leads to the removal of non-specifically bound mEGFP-U2AF65. In the absence of ATP the U1 snRNP dependent association of mEGFP-U2AF65 may still occur, but the detection of this interaction may be masked by the non-specific association of mEGFP-U2AF65 observed on all constructs. The U1 snRNP-dependent association of mEGFP-U2AF65 is not dependent on the pY tract at the 3' SS, ruling out the possibility that U1 snRNP is stimulating the binding of mEGFP-U2AF65 to the pY tract at the 3' SS (Li and Blencowe, 1999; Cote et al., 1995). As a binding site has not been defined for the U1 snRNP-dependent association, it is possible that the observed co-localization does not represent a direct interaction between the mEGFP-U2AF65 and the pre-mRNA, but instead represents an association through protein-protein interactions. With this in mind, one possibility is that the U1

snRNP is stimulating recruitment of the mEGFP-U2AF65 to the pre-mRNA and that this mEGFP-U2AF65 may then be meant to bind the upstream pY tract (Hoffman and Grabowski, 1992). The observation that mEGFP-U2AF65 does not co-localize with the α -U1-Cy5 oligonucleotide suggests that in the absence of pre-mRNA there is no direct association of mEGFP-U2AF65 with the U1 snRNP (Table 5.6). This lack of an interaction suggests that there must be intermediary proteins for this U1 snRNP-dependent association. The constructs used in Section 5.5 lacked an upstream pY tract, but it is possible that the mEGFP-U2AF65 may be recruited, through the interaction of U2AF35 with SR proteins, presumably bound to the exon upstream of the 5' SS (Figure 6.3; Wu and Maniatis, 1993). The dependence on U1 snRNP may arise from the ability of U1 snRNP to stabilise the binding of SR proteins to the pre-mRNA (Jamison et al., 1995; Figure 6.3). Blocking U1 snRNP binding may lead to a loss of stable SR protein association, leading to a loss of mEGFP-U2AF65 association. The dependence of U2AF65 association on SR proteins needs to be tested, assuming that U1 snRNP binds in the absence of SR proteins the dependence of U2AF65 association on SR proteins could be tested by precipitating the SR proteins from an extract containing mEGFP-U2AF65 using high concentrations of $MgCl_2$.

The third mechanism of mEGFP-U2AF65 association with the globin C pre-mRNA appears to be the well-documented association of U2AF65 with the pY tract at the 3' SS (Zamore et al., 1992; Singh et al., 1995; Banerjee et al., 2003; Banerjee et al., 2004). This association of mEGFP-U2AF65 was detected, under A complex conditions and was dependent on the presence of the pY tract at the 3' SS. This suggests that mEGFP-U2AF65 was able to

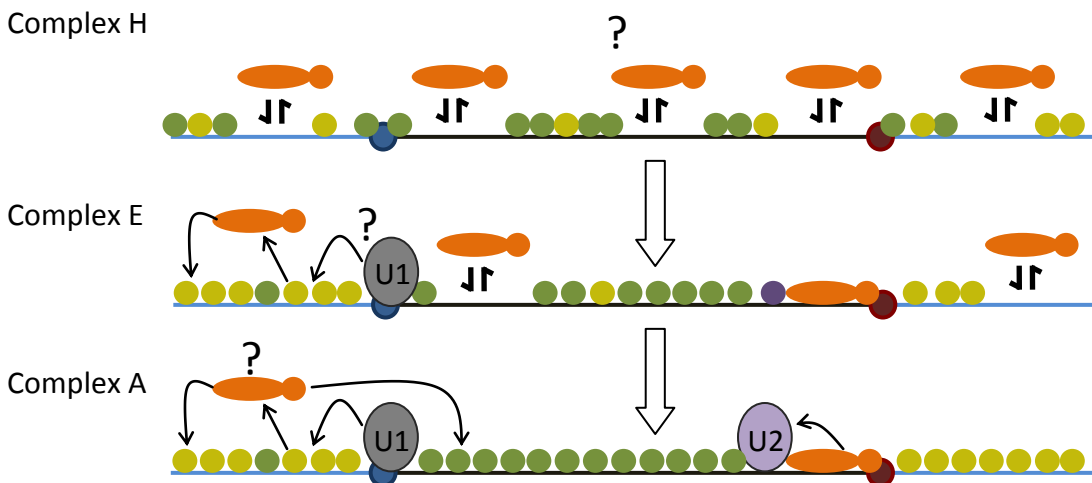


Figure 6.2: Possible mechanisms of U2AF65 binding. Under Complex H conditions the pre-mRNA is bound by a series of non-specific factors, possibly splicing enhancers and splicing inhibitors (yellow and green respectively) as well as an unknown number of U2AF (orange). Under Complex E conditions, the U1 snRNP (grey) specifically associates with the 5' SS and a U2AF associates with the pY tract close to the 3' SS in conjunction with SF1 (purple). Other U2AF may associate at different points along the pre-mRNA; U1 snRNP may stimulate the recruitment of a U2AF. Under Complex A conditions the SF1 at the pY tract is replaced with the U2 snRNP (lilac) and the U1 snRNP promotes the association of a U2AF with the pre-mRNA.

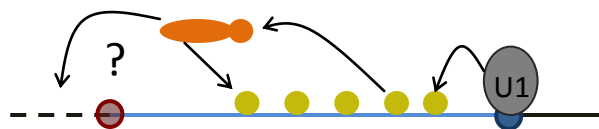


Figure 6.3: Possible role of U1 snRNP in the association of U2AF65 with pre-mRNA. U1 snRNP (grey), bound to the 5' SS (blue) may stimulate binding of factors such as SR proteins (yellow) that in turn stimulate the recruitment of U2AF (Orange), through U2AF35, to the pre-mRNA.

associate with the pY tract. The association was not dependent on U1 snRNP as it still occurred on constructs lacking a 5' SS but containing a pY tract.

Unfortunately the lack of data regarding the number of mEGFP-U2AF65 bound to the globin C pre-mRNAs means it is impossible to assess the number of mEGFP-U2AF65 bound as a result of each of these mechanisms. However the data collected does suggest that an interaction with the pY tract at the 3' SS is not the only way in which U2AF65 can associate with pre-mRNA *in vitro*.

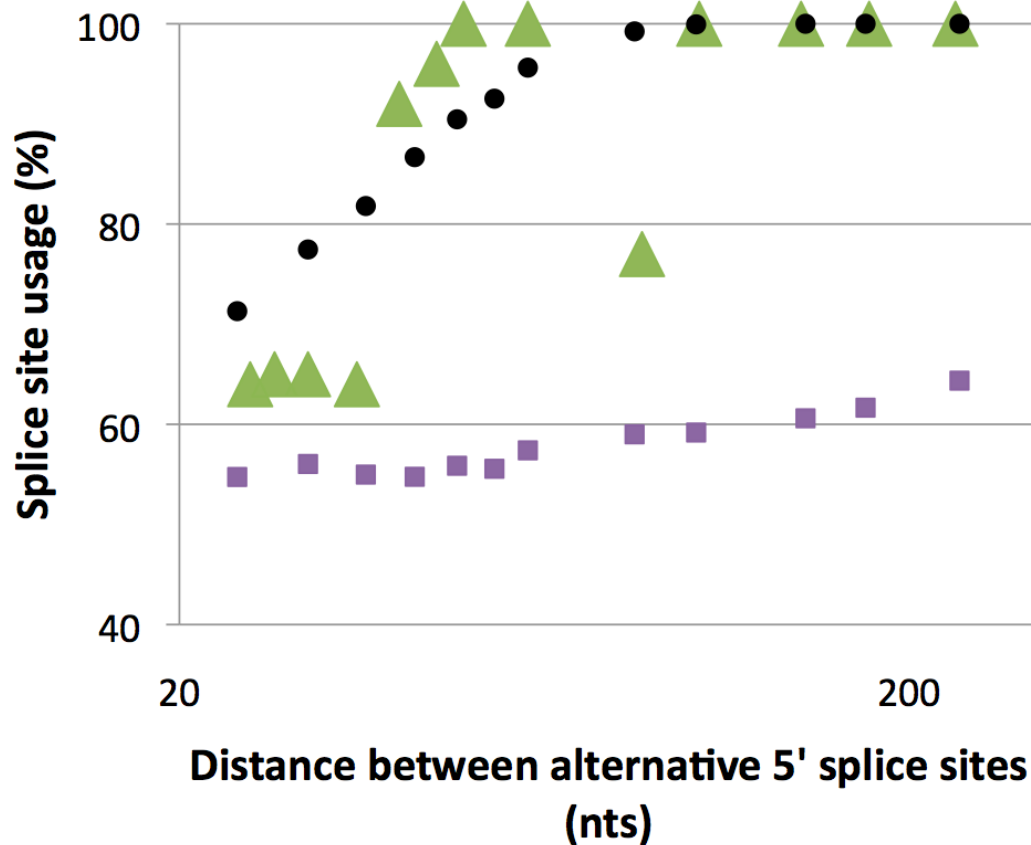
6.6 Concluding remarks

The work I have described lead to the view that the molecules that first recognise and define the splice sites, the U1 snRNP and U2AF, bind promiscuously in E complex but selectively in A complex. This lack of specificity in E complex raises the question of how physiologically relevant it is. Is it an artefact of the *in vitro* splicing systems used throughout the world to study splicing? Why would there be a mechanism whereby U2AF, which is thought to be a specific splicing factor, could bind non-specifically to pre-mRNA?

Interestingly, this illustrates how valuable, techniques such as the TIRF-microscopy used in this study, could be come. These techniques allow analysis of single pre-mRNAs instead of bulk experiments in which conclusions must be drawn from a large population of molecules. This means that instead of focusing on the cross-links that occur between the pY tract and U2AF-65 I have been able to investigate the binding of U2AF65 to the whole transcript. TIRF microscopy could be used to study many aspects of splicing including; the

number of regulatory factors such as SFRS-1 and hnRNP A1 that bind a transcript, is the concentration dependent effect due to increased binding or an increased chance of binding? The point at which proteins associate and are released from the spliceosome could also be analyzed; when does U2AF get removed from the pY tract? In the future, using techniques similar to these used here, it may be possible to study splicing in real-time at the single-pre-mRNA and even to investigate the role of transcription and splicing factors in co-transcriptional splicing. This is a very useful technique that is going to give insights into the inner workings of the spliceosome that no other current technique can.

Appendix 1:



Comparison of observed downstream 5' splice site usage with expected probabilities of collision with the 3' splice site: Green triangles show the percentage of downstream 5' SS use in HeLa cells for globin based constructs with an increasing number of nucleotides between the two 5' SS (Cunningham et al., 1991). Splice site usage at each site is shown, with distance (nts) on a logarithmic scale. Splicing *in vitro* shows the same effect with a transition to exclusively downstream site use with separations above 45 nts. Squares represent the predicted percentage of downstream 5' SS use based upon the chance of collision between the 5' SS and 3' SS determined by a simulation based on the distribution at thermodynamic equilibrium of possible conformations of a homogeneous random coil, modelled as a freely-jointed chain. Circles represent the predicted percentage of downstream 5' SS likewise if the sequence between the splice sites is rigid.

Abbreviations

AG	adenosine-guanine dinucleotide
Bp	base pair
BPS	branch point sequence
BSA	bovine serum albumin
dATP	deoxy adenosine 5' tri-phosphate
dCTP	deoxy cytidine 5' tri-phosphate
DDT	dithiothreitol
dGTP	deoxy guanosine 5' tri-phosphate
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
dTTP	Thymidine 5' tri-phosphate
EDTA	ethylene diamine tetraacetic acid
FBS	fetal bovine serum
HBS	Hepes buffered saline
LB	Luria broth
LMP	low melting point

NE	nuclear extract
NT	nucleotides
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pY tract	polypyrimidine tract
ATP	adenosine 5' tri-phosphate
CTP	cytidine 5' tri-phosphate
GTP	guanosine 5' tri-phosphate
NTP	nucleoside 5' tri-phosphate
UTP	uridine 5' tri-phosphate
SDS	sodium dodecyl sulphate
SF2	splicing factor 2 (SFRS-1)
SFRS-1	splicing factor 2 (SF2)
SOC	super optimal broth with catabolite repression
SSC	standard saline citrate
TBE	tris-borate-EDTA buffer
Tris	tris(hydroxymethyl)aminomethane

5' SS 5' Splice Site

3' SS 3' Splice Site

References

ABBINK, T.E. and BERKHOUT, B., 2008. RNA structure modulates splicing efficiency at the human immunodeficiency virus type 1 major splice donor. *Journal of virology*, **82**(6), pp. 3090-3098.

ABOVICH, N. and ROSBASH, M., 1997. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell*, **89**(3), pp. 403-412.

AITKEN, C.E., MARSHALL, R.A. and PUGLISI, J.D., 2008. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophysical journal*, **94**(5), pp. 1826-1835.

AKERMAN, M. and MANDEL-GUTFREUND, Y., 2006. Alternative splicing regulation at tandem 3' splice sites. *Nucleic acids research*, **34**(1), pp. 23-31.

AKHMEDOV, A.T. and LOPEZ, B.S., 2000. Human 100-kDa homologous DNA-pairing protein is the splicing factor PSF and promotes DNA strand invasion. *Nucleic acids research*, **28**(16), pp. 3022-3030.

ALBERTAZZI, L., AROSIO, D., MARCHETTI, L., RICCI, F. and BELTRAM, F., 2009. Quantitative FRET analysis with the EGFP-mCherry fluorescent protein pair. *Photochemistry and photobiology*, **85**(1), pp. 287-297.

ANDERSON, K. and MOORE, M.J., 1997. Bimolecular exon ligation by the human spliceosome. *Science (New York, N.Y.)*, **276**(5319), pp. 1712-1716.

AST, G., 2004. How did alternative splicing evolve? *Nature reviews.Genetics*, **5**(10), pp. 773-782.

AXELROD, D., BURGHARDT, T.P. and THOMPSON, N.L., 1984. Total internal reflection fluorescence. *Annual Review of Biophysics and Bioengineering*, **13**, pp. 247-268.

BACH, M., KROL, A. and LUHRMANN, R., 1990. Structure-probing of U1 snRNPs gradually depleted of the U1-specific proteins A, C and 70k. Evidence that A interacts differentially with developmentally regulated mouse U1 snRNA variants. *Nucleic acids research*, **18**(3), pp. 449-457.

- BAGSHAW, C.R. and CHERNY, D., 2006. Blinking fluorophores: what do they tell us about protein dynamics? *Biochemical Society transactions*, **34**(Pt 5), pp. 979-982.
- BAI, Y., LEE, D., YU, T. and CHASIN, L.A., 1999. Control of 3' splice site choice in vivo by ASF/SF2 and hnRNP A1. *Nucleic acids research*, **27**(4), pp. 1126-1134.
- BANERJEE, H., RAHN, A., DAVIS, W. and SINGH, R., 2003. Sex lethal and U2 small nuclear ribonucleoprotein auxiliary factor (U2AF65) recognize polypyrimidine tracts using multiple modes of binding. *RNA (New York, N.Y.)*, **9**(1), pp. 88-99.
- BANERJEE, H., RAHN, A., GAWANDE, B., GUTH, S., VALCARCEL, J. and SINGH, R., 2004. The conserved RNA recognition motif 3 of U2 snRNA auxiliary factor (U2AF 65) is essential in vivo but dispensable for activity in vitro. *RNA (New York, N.Y.)*, **10**(2), pp. 240-253.
- BARASH, Y., CALARCO, J.A., GAO, W., PAN, Q., WANG, X., SHAI, O., BLENCOWE, B.J. and FREY, B.J., 2010. Deciphering the splicing code. *Nature*, **465**(7294), pp. 53-59.
- BEHZADNIA, N., GOLAS, M.M., HARTMUTH, K., SANDER, B., KASTNER, B., DECKERT, J., DUBE, P., WILL, C.L., URLAUB, H., STARK, H. and LUHRMANN, R., 2007. Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. *The EMBO journal*, **26**(6), pp. 1737-1748.
- BEN-DOV, C., HARTMANN, B., LUNDGREN, J. and VALCARCEL, J., 2008. Genome-wide analysis of alternative pre-mRNA splicing. *The Journal of biological chemistry*, **283**(3), pp. 1229-1233.
- BERGET, S.M., 1995. Exon recognition in vertebrate splicing. *The Journal of biological chemistry*, **270**(6), pp. 2411-2414.
- BERGLUND, J.A., ABOVICH, N. and ROSBASH, M., 1998. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes & development*, **12**(6), pp. 858-867.
- BERGLUND, J.A., CHUA, K., ABOVICH, N., REED, R. and ROSBASH, M., 1997. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell*, **89**(5), pp. 781-787.
- BESSONOV, S., ANOKHINA, M., KRASAUSKAS, A., GOLAS, M.M., SANDER, B., WILL, C.L., URLAUB, H., STARK, H. and LUHRMANN, R., 2010. Characterization of purified human Bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. *RNA (New York, N.Y.)*, **16**(12), pp. 2384-2403.

BLACK, D.L., 1998. Splicing in the inner ear: a familiar tune, but what are the instruments? *Neuron*, **20**(2), pp. 165-168.

BLACK, D.L., CHABOT, B. and STEITZ, J.A., 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. *Cell*, **42**(3), pp. 737-750.

BLACK, D.L. and STEITZ, J.A., 1986. Pre-mRNA splicing in vitro requires intact U4/U6 small nuclear ribonucleoprotein. *Cell*, **46**(5), pp. 697-704.

BLANCHETTE, M. and CHABOT, B., 1999. Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. *The EMBO journal*, **18**(7), pp. 1939-1952.

BLANCHETTE, M. and CHABOT, B., 1997. A highly stable duplex structure sequesters the 5' splice site region of hnRNP A1 alternative exon 7B. *RNA (New York, N.Y.)*, **3**(4), pp. 405-419.

BLEICHERT, F. and BASERGA, S.J., 2007. The long unwinding road of RNA helicases. *Molecular cell*, **27**(3), pp. 339-352.

BLENCOWE, B.J., ISSNER, R., NICKERSON, J.A. and SHARP, P.A., 1998. A coactivator of pre-mRNA splicing. *Genes & development*, **12**(7), pp. 996-1009.

BLENCOWE, B.J., SPROAT, B.S., RYDER, U., BARABINO, S. and LAMOND, A.I., 1989. Antisense probing of the human U4/U6 snRNP with biotinylated 2'-OMe RNA oligonucleotides. *Cell*, **59**(3), pp. 531-539.

BONNAL, S., MARTINEZ, C., FORCH, P., BACHI, A., WILM, M. and VALCARCEL, J., 2008. RBM5/Luca-15/H37 regulates Fas alternative splice site pairing after exon definition. *Molecular cell*, **32**(1), pp. 81-95.

BOUCK, J., FU, X.D., SKALKA, A.M. and KATZ, R.A., 1998. Role of the constitutive splicing factors U2AF65 and SAP49 in suboptimal RNA splicing of novel retroviral mutants. *The Journal of biological chemistry*, **273**(24), pp. 15169-15176.

BOUKIS, L.A., LIU, N., FURUYAMA, S. and BRUZIK, J.P., 2004. Ser/Arg-rich protein-mediated communication between U1 and U2 small nuclear ribonucleoprotein particles. *The Journal of biological chemistry*, **279**(28), pp. 29647-29653.

BOURGEOIS, C.F., POPIELARZ, M., HILDWEIN, G. and STEVENIN, J., 1999. Identification of a bidirectional splicing enhancer: differential involvement of SR proteins in 5' or 3' splice site activation. *Molecular and cellular biology*, **19**(11), pp. 7347-7356.

- BRANLANT, C., KROL, A., EBEL, J.P., GALLINARO, H., LAZAR, E. and JACOB, M., 1981. The conformation of chicken, rat and human U1A RNAs in solution. *Nucleic acids research*, **9**(4), pp. 841-858.
- BRANLANT, C., KROL, A., EBEL, J.P., LAZAR, E., HAENDLER, B. and JACOB, M., 1982. U2 RNA shares a structural domain with U1, U4, and U5 RNAs. *The EMBO journal*, **1**(10), pp. 1259-1265.
- BROSI, R., GRONING, K., BEHRENS, S.E., LUHRMANN, R. and KRAMER, A., 1993a. Interaction of mammalian splicing factor SF3a with U2 snRNP and relation of its 60-kD subunit to yeast PRP9. *Science (New York, N.Y.)*, **262**(5130), pp. 102-105.
- BROSI, R., HAURI, H.P. and KRAMER, A., 1993b. Separation of splicing factor SF3 into two components and purification of SF3a activity. *The Journal of biological chemistry*, **268**(23), pp. 17640-17646.
- BURATTI, E., BARALLE, M., DE CONTI, L., BARALLE, D., ROMANO, M., AYALA, Y.M. and BARALLE, F.E., 2004. hnRNP H binding at the 5' splice site correlates with the pathological effect of two intronic mutations in the NF-1 and TSHbeta genes. *Nucleic acids research*, **32**(14), pp. 4224-4236.
- CACERES, J.F. and KRAINER, A.R., 1993. Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. *The EMBO journal*, **12**(12), pp. 4715-4726.
- CACERES, J.F., MISTELI, T., SCREATON, G.R., SPECTOR, D.L. and KRAINER, A.R., 1997. Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *The Journal of cell biology*, **138**(2), pp. 225-238.
- CACERES, J.F., STAMM, S., HELFMAN, D.M. and KRAINER, A.R., 1994. Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science (New York, N.Y.)*, **265**(5179), pp. 1706-1709.
- CAPUTI, M. and ZAHLER, A.M., 2002. SR proteins and hnRNP H regulate the splicing of the HIV-1 tev-specific exon 6D. *The EMBO journal*, **21**(4), pp. 845-855.
- CASTLE, J.C., ARMOUR, C.D., LOWER, M., HAYNOR, D., BIERY, M., BOUZEK, H., CHEN, R., JACKSON, S., JOHNSON, J.M., ROHL, C.A. and RAYMOND, C.K., 2010. Digital genome-wide ncRNA expression, including SnoRNAs, across 11 human tissues using polyA-neutral amplification. *PloS one*, **5**(7), pp. e11779.
- CASTLE, J.C., ZHANG, C., SHAH, J.K., KULKARNI, A.V., KALSOTRA, A., COOPER, T.A. and JOHNSON, J.M., 2008. Expression of 24,426 human alternative splicing events and predicted cis regulation in 48 tissues and cell lines. *Nature genetics*, **40**(12), pp. 1416-1425.

CHABOT, B., BLACK, D.L., LEMASTER, D.M. and STEITZ, J.A., 1985. The 3' splice site of pre-messenger RNA is recognized by a small nuclear ribonucleoprotein. *Science (New York, N.Y.)*, **230**(4732), pp. 1344-1349.

CHABOT, B., BLANCHETTE, M., LAPIERRE, I. and LA BRANCHE, H., 1997. An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1. *Molecular and cellular biology*, **17**(4), pp. 1776-1786.

CHANG, D.D. and SHARP, P.A., 1989. Regulation by HIV Rev depends upon recognition of splice sites. *Cell*, **59**(5), pp. 789-795.

CHAOUKI, A.S. and SALZ, H.K., 2006. Drosophila SPF45: a bifunctional protein with roles in both splicing and DNA repair. *PLoS genetics*, **2**(12), pp. e178.

CHATTORAJ, M., KING, B.A., BUBLITZ, G.U. and BOXER, S.G., 1996. Ultra-fast excited state dynamics in green fluorescent protein: multiple states and proton transfer. *Proceedings of the National Academy of Sciences of the United States of America*, **93**(16), pp. 8362-8367.

CHEN, C., ZHAO, X., KIERZEK, R. and YU, Y.T., 2010. A flexible RNA backbone within the polypyrimidine tract is required for U2AF65 binding and pre-mRNA splicing in vivo. *Molecular and cellular biology*, **30**(17), pp. 4108-4119.

CHEN, J.Y., STANDS, L., STALEY, J.P., JACKUPS, R.R., JR, LATUS, L.J. and CHANG, T.H., 2001. Specific alterations of U1-C protein or U1 small nuclear RNA can eliminate the requirement of Prp28p, an essential DEAD box splicing factor. *Molecular cell*, **7**(1), pp. 227-232.

CHEN, S., ANDERSON, K. and MOORE, M.J., 2000. Evidence for a linear search in bimolecular 3' splice site AG selection. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(2), pp. 593-598.

CHERNY, D., GOODING, C., EPERON, G.E., COELHO, M.B., BAGSHAW, C.R., SMITH, C.W. and EPERON, I.C., 2010. Stoichiometry of a regulatory splicing complex revealed by single-molecule analyses. *The EMBO journal*, **29**(13), pp. 2161-2172.

CHIARA, M.D., PALANDJIAN, L., FELD KRAMER, R. and REED, R., 1997. Evidence that U5 snRNP recognizes the 3' splice site for catalytic step II in mammals. *The EMBO journal*, **16**(15), pp. 4746-4759.

CHO, S., HOANG, A., SINHA, R., ZHONG, X.Y., FU, X.D., KRAINER, A.R. and GHOSH, G., 2011. Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K snRNP protein determines early spliceosome assembly. *Proceedings of the National Academy of Sciences of the United States of America*, . doi: 10.1073/pnas.1017700108

- CHRISTOFORI, G., FRENDEWEY, D. and KELLER, W., 1987. Two spliceosomes can form simultaneously and independently on synthetic double-intron messenger RNA precursors. *The EMBO journal*, **6**(6), pp. 1747-1755.
- CHUA, K. and REED, R., 2001. An upstream AG determines whether a downstream AG is selected during catalytic step II of splicing. *Molecular and cellular biology*, **21**(5), pp. 1509-1514.
- CHUA, K. and REED, R., 1999. The RNA splicing factor hSlu7 is required for correct 3' splice-site choice. *Nature*, **402**(6758), pp. 207-210.
- CHUSAINOW, J., AJUH, P.M., TRINKLE-MULCAHY, L., SLEEMAN, J.E., ELLENBERG, J. and LAMOND, A.I., 2005. FRET analyses of the U2AF complex localize the U2AF35/U2AF65 interaction in vivo and reveal a novel self-interaction of U2AF35. *RNA (New York, N.Y.)*, **11**(8), pp. 1201-1214.
- COHEN-KRAUSZ, S., SPERLING, R. and SPERLING, J., 2007. Exploring the architecture of the intact supraspliceosome using electron microscopy. *Journal of Molecular Biology*, **368**(2), pp. 319-327.
- CONIBEAR, P.B. and BAGSHAW, C.R., 2000. A comparison of optical geometries for combined flash photolysis and total internal reflection fluorescence microscopy. *Journal of microscopy*, **200**(Pt 3), pp. 218-229.
- COOLIDGE, C.J., SEELY, R.J. and PATTON, J.G., 1997. Functional analysis of the polypyrimidine tract in pre-mRNA splicing. *Nucleic acids research*, **25**(4), pp. 888-896.
- CORTES, J.J., SONTHEIMER, E.J., SEIWERT, S.D. and STEITZ, J.A., 1993. Mutations in the conserved loop of human U5 snRNA generate use of novel cryptic 5' splice sites in vivo. *The EMBO journal*, **12**(13), pp. 5181-5189.
- COTE, J., SIMARD, M.J. and CHABOT, B., 1999. An element in the 5' common exon of the NCAM alternative splicing unit interacts with SR proteins and modulates 5' splice site selection. *Nucleic acids research*, **27**(12), pp. 2529-2537.
- COTTEN, M., OBERHAUSER, B., BRUNAR, H., HOLZNER, A., ISSAKIDES, G., NOE, C.R., SCHAFFNER, G., WAGNER, E. and BIRNSTIEL, M.L., 1991. 2'-O-methyl, 2'-O-ethyl oligoribonucleotides and phosphorothioate oligodeoxyribonucleotides as inhibitors of the in vitro U7 snRNP-dependent mRNA processing event. *Nucleic acids research*, **19**(10), pp. 2629-2635.
- CRAWFORD, D.J., HOSKINS, A.A., FRIEDMAN, L.J., GELLES, J. and MOORE, M.J., 2008. Visualizing the splicing of single pre-mRNA molecules in whole cell extract. *RNA (New York, N.Y.)*, **14**(1), pp. 170-179.

- CREEMERS, T.M., LOCK, A.J., SUBRAMANIAM, V., JOVIN, T.M. and VOLKER, S., 1999. Three photoconvertible forms of green fluorescent protein identified by spectral hole-burning. *Nature structural biology*, **6**(6), pp. 557-560.
- CRISPINO, J.D., BLENCOWE, B.J. and SHARP, P.A., 1994. Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science (New York, N.Y.)*, **265**(5180), pp. 1866-1869.
- CUBITT, A.B., HEIM, R., ADAMS, S.R., BOYD, A.E., GROSS, L.A. and TSIEN, R.Y., 1995. Understanding, improving and using green fluorescent proteins. *Trends in biochemical sciences*, **20**(11), pp. 448-455.
- CUNNINGHAM, S.A., ELSE, A.J., POTTER, B.V. and EPERON, I.C., 1991. Influences of separation and adjacent sequences on the use of alternative 5' splice sites. *Journal of Molecular Biology*, **217**(2), pp. 265-281.
- DAS, R. and REED, R., 1999. Resolution of the mammalian E complex and the ATP-dependent spliceosomal complexes on native agarose mini-gels. *RNA (New York, N.Y.)*, **5**(11), pp. 1504-1508.
- DAS, R., ZHOU, Z. and REED, R., 2000. Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. *Molecular cell*, **5**(5), pp. 779-787.
- DAVIES, M.J., 2004. Reactive species formed on proteins exposed to singlet oxygen. *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology*, **3**(1), pp. 17-25.
- DAYIE, K.T. and PADGETT, R.A., 2008. A glimpse into the active site of a group II intron and maybe the spliceosome, too. *RNA (New York, N.Y.)*, **14**(9), pp. 1697-1703.
- DECKERT, J., HARTMUTH, K., BOEHRINGER, D., BEHZADNIA, N., WILL, C.L., KASTNER, B., STARK, H., URLAUB, H. and LUHRMANN, R., 2006. Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. *Molecular and cellular biology*, **26**(14), pp. 5528-5543.
- DEUTSCH, M. and LONG, M., 1999. Intron-exon structures of eukaryotic model organisms. *Nucleic acids research*, **27**(15), pp. 3219-3228.
- DICKSON, R.M., CUBITT, A.B., TSIEN, R.Y. and MOERNER, W.E., 1997. On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature*, **388**(6640), pp. 355-358.

DONEV, R.M., DONEVA, T.A., BOWEN, W.R. and SHEER, D., 2002. HnRNP-A1 binds directly to double-stranded DNA in vitro within a 36 bp sequence. *Molecular and cellular biochemistry*, **233**(1-2), pp. 181-185.

DONIS-KELLER, H., 1979. Site specific enzymatic cleavage of RNA. *Nucleic acids research*, **7**(1), pp. 179-192.

DONMEZ, G., HARTMUTH, K., KASTNER, B., WILL, C.L. and LUHRMANN, R., 2007. The 5' end of U2 snRNA is in close proximity to U1 and functional sites of the pre-mRNA in early spliceosomal complexes. *Molecular cell*, **25**(3), pp. 399-411.

DONMEZ, G., HARTMUTH, K. and LUHRMANN, R., 2004. Modified nucleotides at the 5' end of human U2 snRNA are required for spliceosomal E-complex formation. *RNA (New York, N.Y.)*, **10**(12), pp. 1925-1933.

DUBRIDGE, R.B., TANG, P., HSIA, H.C., LEONG, P.M., MILLER, J.H. and CALOS, M.P., 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Molecular and cellular biology*, **7**(1), pp. 379-387.

DU, H. and ROSBASH, M., 2002. The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing. *Nature*, **419**(6902), pp. 86-90.

ELDRIDGE, A.G., LI, Y., SHARP, P.A. and BLENCOWE, B.J., 1999. The SRm160/300 splicing coactivator is required for exon-enhancer function. *Proceedings of the National Academy of Sciences of the United States of America*, **96**(11), pp. 6125-6130.

EPERON, I.C., IRELAND, D.C., SMITH, R.A., MAYEDA, A. and KRAINER, A.R., 1993. Pathways for selection of 5' splice sites by U1 snRNPs and SF2/ASF. *The EMBO journal*, **12**(9), pp. 3607-3617.

EPERON, I.C., MAKAROVA, O.V., MAYEDA, A., MUNROE, S.H., CACERES, J.F., HAYWARD, D.G. and KRAINER, A.R., 2000. Selection of alternative 5' splice sites: role of U1 snRNP and models for the antagonistic effects of SF2/ASF and hnRNP A1. *Molecular and cellular biology*, **20**(22), pp. 8303-8318.

EPERON, L.P., ESTIBEIRO, J.P. and EPERON, I.C., 1986. The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA. *Nature*, **324**(6094), pp. 280-282.

EPERON, L.P., GRAHAM, I.R., GRIFFITHS, A.D. and EPERON, I.C., 1988. Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell*, **54**(3), pp. 393-401.

- FISETTE, J.F., TOUTANT, J., DUGRE-BRISSON, S., DESGROSEILLERS, L. and CHABOT, B., 2010. hnRNP A1 and hnRNP H can collaborate to modulate 5' splice site selection. *RNA (New York, N.Y.)*, **16**(1), pp. 228-238.
- FISET, S. and CHABOT, B., 2001. hnRNP A1 may interact simultaneously with telomeric DNA and the human telomerase RNA in vitro. *Nucleic acids research*, **29**(11), pp. 2268-2275.
- FLECKNER, J., ZHANG, M., VALCARCEL, J. and GREEN, M.R., 1997. U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes & development*, **11**(14), pp. 1864-1872.
- FORCH, P., MERENDINO, L., MARTINEZ, C. and VALCARCEL, J., 2003. U2 small nuclear ribonucleoprotein particle (snRNP) auxiliary factor of 65 kDa, U2AF65, can promote U1 snRNP recruitment to 5' splice sites. *The Biochemical journal*, **372**(Pt 1), pp. 235-240.
- FORCH, P., MERENDINO, L., MARTINEZ, C. and VALCARCEL, J., 2001. Modulation of msl-2 5' splice site recognition by Sex-lethal. *RNA (New York, N.Y.)*, **7**(9), pp. 1185-1191.
- FORCH, P., PUIG, O., KEDERSHA, N., MARTINEZ, C., GRANNEMAN, S., SERAPHIN, B., ANDERSON, P. and VALCARCEL, J., 2000. The apoptosis-promoting factor TIA-1 is a regulator of alternative pre-mRNA splicing. *Molecular cell*, **6**(5), pp. 1089-1098.
- FORCH, P., PUIG, O., MARTINEZ, C., SERAPHIN, B. and VALCARCEL, J., 2002. The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5' splice sites. *The EMBO journal*, **21**(24), pp. 6882-6892.
- FOX-WALSH, K.L., DOU, Y., LAM, B.J., HUNG, S.P., BALDI, P.F. and HERTEL, K.J., 2005. The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(45), pp. 16176-16181.
- FRENDEWEY, D. and KELLER, W., 1985. Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. *Cell*, **42**(1), pp. 355-367.
- FREUND, M., ASANG, C., KAMMLER, S., KONERMANN, C., KRUMMHEUER, J., HIPPEL, M., MEYER, I., GIERLING, W., THEISS, S., PREUSS, T., SCHINDLER, D., KJEMS, J. and SCHAAL, H., 2003. A novel approach to describe a U1 snRNA binding site. *Nucleic acids research*, **31**(23), pp. 6963-6975.
- FU, X.D., MAYEDA, A., MANIATIS, T. and KRAINER, A.R., 1992. General splicing factors SF2 and SC35 have equivalent activities in vitro, and both affect alternative 5' and 3' splice site selection. *Proceedings of the National Academy of Sciences of the United States of America*, **89**(23), pp. 11224-11228.

- FUKUMURA, K., TANIGUCHI, I., SAKAMOTO, H., OHNO, M. and INOUE, K., 2009. U1-independent pre-mRNA splicing contributes to the regulation of alternative splicing. *Nucleic acids research*, **37**(6), pp. 1907-1914.
- FUNATSU, T., HARADA, Y., TOKUNAGA, M., SAITO, K. and YANAGIDA, T., 1995. Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature*, **374**(6522), pp. 555-559.
- GAL-MARK, N., SCHWARTZ, S., RAM, O., EYRAS, E. and AST, G., 2009. The pivotal roles of TIA proteins in 5' splice-site selection of alu exons and across evolution. *PLoS genetics*, **5**(11), pp. e1000717.
- GAMA-CARVALHO, M., KRAUSS, R.D., CHIANG, L., VALCARCEL, J., GREEN, M.R. and CARMO-FONSECA, M., 1997. Targeting of U2AF65 to sites of active splicing in the nucleus. *The Journal of cell biology*, **137**(5), pp. 975-987.
- GAO, K., MASUDA, A., MATSUURA, T. and OHNO, K., 2008. Human branch point consensus sequence is yUnAy. *Nucleic acids research*, **36**(7), pp. 2257-2267.
- GE, H. and MANLEY, J.L., 1990. A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. *Cell*, **62**(1), pp. 25-34.
- GOLOVANOV, A.P., HAUTBERGUE, G.M., WILSON, S.A. and LIAN, L.Y., 2004. A simple method for improving protein solubility and long-term stability. *Journal of the American Chemical Society*, **126**(29), pp. 8933-8939.
- GOODING, C., CLARK, F., WOLLERTON, M.C., GRELLSCHEID, S.N., GROOM, H. and SMITH, C.W., 2006. A class of human exons with predicted distant branch points revealed by analysis of AG dinucleotide exclusion zones. *Genome biology*, **7**(1), pp. R1.
- GOODING, C., ROBERTS, G.C., MOREAU, G., NADAL-GINARD, B. and SMITH, C.W., 1994. Smooth muscle-specific switching of alpha-tropomyosin mutually exclusive exon selection by specific inhibition of the strong default exon. *The EMBO journal*, **13**(16), pp. 3861-3872.
- GORNEMANN, J., KOTOVIC, K.M., HUJER, K. and NEUGEBAUER, K.M., 2005. Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Molecular cell*, **19**(1), pp. 53-63.
- GOZANI, O., FELD, R. and REED, R., 1996. Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes & development*, **10**(2), pp. 233-243.

- GOZANI, O., PATTON, J.G. and REED, R., 1994. A novel set of spliceosome-associated proteins and the essential splicing factor PSF bind stably to pre-mRNA prior to catalytic step II of the splicing reaction. *The EMBO journal*, **13**(14), pp. 3356-3367.
- GOZANI, O., POTASHKIN, J. and REED, R., 1998. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Molecular and cellular biology*, **18**(8), pp. 4752-4760.
- GRABOWSKI, P.J., PADGETT, R.A. and SHARP, P.A., 1984. Messenger RNA splicing in vitro: an excised intervening sequence and a potential intermediate. *Cell*, **37**(2), pp. 415-427.
- GRAVELEY, B.R., 2000. Sorting out the complexity of SR protein functions. *RNA (New York, N.Y.)*, **6**(9), pp. 1197-1211.
- GRAVELEY, B.R., HERTEL, K.J. and MANIATIS, T., 2001. The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA (New York, N.Y.)*, **7**(6), pp. 806-818.
- GRAVELEY, B.R. and MANIATIS, T., 1998. Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Molecular cell*, **1**(5), pp. 765-771.
- GREENBAUM, L., ROTHMANN, C., LAVIE, R. and MALIK, Z., 2000. Green fluorescent protein photobleaching: a model for protein damage by endogenous and exogenous singlet oxygen. *Biological chemistry*, **381**(12), pp. 1251-1258.
- GUPTA, A., JENKINS, J.L. and KIELKOPF, C.L., 2011. RNA induces conformational changes in the SF1/U2AF65 splicing factor complex. *Journal of Molecular Biology*, **405**(5), pp. 1128-1138.
- GUTH, S., MARTINEZ, C., GAUR, R.K. and VALCARCEL, J., 1999. Evidence for substrate-specific requirement of the splicing factor U2AF(35) and for its function after polypyrimidine tract recognition by U2AF(65). *Molecular and cellular biology*, **19**(12), pp. 8263-8271.
- GUTH, S., TANGE, T.O., KELLENBERGER, E. and VALCARCEL, J., 2001. Dual function for U2AF(35) in AG-dependent pre-mRNA splicing. *Molecular and cellular biology*, **21**(22), pp. 7673-7681.
- HARTMUTH, K., URLAUB, H., VORNLOCHER, H.P., WILL, C.L., GENTZEL, M., WILM, M. and LUHRMANN, R., 2002. Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *Proceedings of the National Academy of Sciences of the United States of America*, **99**(26), pp. 16719-16724.

- HEIM, R., PRASHER, D.C. and TSIEN, R.Y., 1994. Wavelength mutations and posttranslational autooxidation of green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, **91**(26), pp. 12501-12504.
- HEINRICHS, V., BACH, M., WINKELMANN, G. and LUHRMANN, R., 1990. U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5' splice site. *Science (New York, N.Y.)*, **247**(4938), pp. 69-72.
- HELFMAN, D.M. and RICCI, W.M., 1989. Branch point selection in alternative splicing of tropomyosin pre-mRNAs. *Nucleic acids research*, **17**(14), pp. 5633-5650.
- HENSCHIED, K.L., VOELKER, R.B. and BERGLUND, J.A., 2008. Alternative modes of binding by U2AF65 at the polypyrimidine tract. *Biochemistry*, **47**(1), pp. 449-459.
- HICKS, M.J., MUELLER, W.F., SHEPARD, P.J. and HERTEL, K.J., 2010. Competing upstream 5' splice sites enhance the rate of proximal splicing. *Molecular and cellular biology*, **30**(8), pp. 1878-1886.
- HILLER, M., HUSE, K., SZAFRANSKI, K., JAHN, N., HAMPE, J., SCHREIBER, S., BACKOFEN, R. and PLATZER, M., 2004. Widespread occurrence of alternative splicing at NAGNAG acceptors contributes to proteome plasticity. *Nature genetics*, **36**(12), pp. 1255-1257.
- HOCHLEITNER, E.O., KASTNER, B., FROHLICH, T., SCHMIDT, A., LUHRMANN, R., ARNOLD, G. and LOTTSPREICH, F., 2005. Protein stoichiometry of a multiprotein complex, the human spliceosomal U1 small nuclear ribonucleoprotein: absolute quantification using isotope-coded tags and mass spectrometry. *The Journal of biological chemistry*, **280**(4), pp. 2536-2542.
- HOFFMAN, B.E. and GRABOWSKI, P.J., 1992. U1 snRNP targets an essential splicing factor, U2AF65, to the 3' splice site by a network of interactions spanning the exon. *Genes & development*, **6**(12B), pp. 2554-2568.
- HONG, W., BENNETT, M., XIAO, Y., FELD KRAMER, R., WANG, C. and REED, R., 1997. Association of U2 snRNP with the spliceosomal complex E. *Nucleic acids research*, **25**(2), pp. 354-361.
- HOSKINS, A.A., FRIEDMAN, L.J., GALLAGHER, S.S., CRAWFORD, D.J., ANDERSON, E.G., WOMBACHER, R., RAMIREZ, N., CORNISH, V.W., GELLES, J. and MOORE, M.J., 2011. Ordered and dynamic assembly of single spliceosomes. *Science (New York, N.Y.)*, **331**(6022), pp. 1289-1295.
- HOUSE, A.E. and LYNCH, K.W., 2006. An exonic splicing silencer represses spliceosome assembly after ATP-dependent exon recognition. *Nature structural & molecular biology*, **13**(10), pp. 937-944.

HUANG, Y. and CARMICHAEL, G.G., 1996. A suboptimal 5' splice site is a cis-acting determinant of nuclear export of polyomavirus late mRNAs. *Molecular and cellular biology*, **16**(11), pp. 6046-6054.

HURANOVA, M., IVANI, I., BENDA, A., POSER, I., BRODY, Y., HOF, M., SHAV-TAL, Y., NEUGEBAUER, K.M. and STANEK, D., 2010. The differential interaction of snRNPs with pre-mRNA reveals splicing kinetics in living cells. *The Journal of cell biology*, **191**(1), pp. 75-86.

HWANG, D.Y. and COHEN, J.B., 1996. U1 snRNA promotes the selection of nearby 5' splice sites by U6 snRNA in mammalian cells. *Genes & development*, **10**(3), pp. 338-350.

IBRAHIM, E.C., SCHAAL, T.D., HERTEL, K.J., REED, R. and MANIATIS, T., 2005. Serine/arginine-rich protein-dependent suppression of exon skipping by exonic splicing enhancers. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(14), pp. 5002-5007.

INOUE, K., HOSHIJIMA, K., SAKAMOTO, H. and SHIMURA, Y., 1990. Binding of the Drosophila sex-lethal gene product to the alternative splice site of transformer primary transcript. *Nature*, **344**(6265), pp. 461-463.

IZQUIERDO, J.M., MAJOS, N., BONNAL, S., MARTINEZ, C., CASTELO, R., GUIGO, R., BILBAO, D. and VALCARCEL, J., 2005. Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. *Molecular cell*, **19**(4), pp. 475-484.

JACQUENET, S., MEREAU, A., BILODEAU, P.S., DAMIER, L., STOLTZFUS, C.M. and BRANLANT, C., 2001. A second exon splicing silencer within human immunodeficiency virus type 1 tat exon 2 represses splicing of Tat mRNA and binds protein hnRNP H. *The Journal of biological chemistry*, **276**(44), pp. 40464-40475.

JAMISON, S.F., CROW, A. and GARCIA-BLANCO, M.A., 1992. The spliceosome assembly pathway in mammalian extracts. *Molecular and cellular biology*, **12**(10), pp. 4279-4287.

JAMISON, S.F. and GARCIA-BLANCO, M.A., 1992. An ATP-independent U2 small nuclear ribonucleoprotein particle/precursor mRNA complex requires both splice sites and the polypyrimidine tract. *Proceedings of the National Academy of Sciences of the United States of America*, **89**(12), pp. 5482-5486.

JAMISON, S.F., PASMAN, Z., WANG, J., WILL, C., LUHRMANN, R., MANLEY, J.L. and GARCIA-BLANCO, M.A., 1995. U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements. *Nucleic acids research*, **23**(16), pp. 3260-3267.

JENKINS, J.L., SHEN, H., GREEN, M.R. and KIELKOPF, C.L., 2008. Solution conformation and thermodynamic characteristics of RNA binding by the splicing factor U2AF65. *The Journal of biological chemistry*, **283**(48), pp. 33641-33649.

JIANG, Z., COTE, J., KWON, J.M., GOATE, A.M. and WU, J.Y., 2000. Aberrant splicing of tau pre-mRNA caused by intronic mutations associated with the inherited dementia frontotemporal dementia with parkinsonism linked to chromosome 17. *Molecular and cellular biology*, **20**(11), pp. 4036-4048.

JURICA, M.S., LICKLIDER, L.J., GYGI, S.R., GRIGORIEFF, N. and MOORE, M.J., 2002. Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA (New York, N.Y.)*, **8**(4), pp. 426-439.

JURICA, M.S., SOUSA, D., MOORE, M.J. and GRIGORIEFF, N., 2004. Three-dimensional structure of C complex spliceosomes by electron microscopy. *Nature structural & molecular biology*, **11**(3), pp. 265-269.

KAMBACH, C. and MATTAJ, I.W., 1992. Intracellular distribution of the U1A protein depends on active transport and nuclear binding to U1 snRNA. *The Journal of cell biology*, **118**(1), pp. 11-21.

KAMBACH, C., WALKE, S., YOUNG, R., AVIS, J.M., DE LA FORTELLE, E., RAKER, V.A., LUHRMANN, R., LI, J., NAGAI, K., 1999. Crystal structures of two SM protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell*, **96**, pp. 375-387.

KAN, J.L. and GREEN, M.R., 1999. Pre-mRNA splicing of IgM exons M1 and M2 is directed by a juxtaposed splicing enhancer and inhibitor. *Genes & development*, **13**(4), pp. 462-471.

KANOPKA, A., MUHLEMANN, O. and AKUSJARVI, G., 1996. Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature*, **381**(6582), pp. 535-538.

KELLENBERGER, E., STIER, G. and SATTLER, M., 2002. Induced folding of the U2AF35 RRM upon binding to U2AF65. *FEBS letters*, **528**(1-3), pp. 171-176.

KELLER, E.B. and NOON, W.A., 1984. Intron splicing: a conserved internal signal in introns of animal pre-mRNAs. *Proceedings of the National Academy of Sciences of the United States of America*, **81**(23), pp. 7417-7420.

KENT, O.A. and MACMILLAN, A.M., 2002. Early organization of pre-mRNA during spliceosome assembly. *Nature structural biology*, **9**(8), pp. 576-581.

KENT, O.A., REAYI, A., FOONG, L., CHILIBECK, K.A. and MACMILLAN, A.M., 2003. Structuring of the 3' splice site by U2AF65. *The Journal of biological chemistry*, **278**(50), pp. 50572-50577.

KENT, O.A., RITCHIE, D.B. and MACMILLAN, A.M., 2005. Characterization of a U2AF-independent commitment complex (E') in the mammalian spliceosome assembly pathway. *Molecular and cellular biology*, **25**(1), pp. 233-240.

- KIELKOPF, C.L., RODIONOVA, N.A., GREEN, M.R. and BURLEY, S.K., 2001. A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. *Cell*, **106**(5), pp. 595-605.
- KISTLER, A.L. and GUTHRIE, C., 2001. Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase. *Genes & development*, **15**(1), pp. 42-49.
- KLEIN GUNNEWIEK, J.M., HUSSEIN, R.I., VAN AARSEN, Y., PALACIOS, D., DE JONG, R., VAN VENROOIJ, W.J. and GUNDERSON, S.I., 2000. Fourteen residues of the U1 snRNP-specific U1A protein are required for homodimerization, cooperative RNA binding, and inhibition of polyadenylation. *Molecular and cellular biology*, **20**(6), pp. 2209-2217.
- KOHTZ, J.D., JAMISON, S.F., WILL, C.L., ZUO, P., LUHRMANN, R., GARCIA-BLANCO, M.A. and MANLEY, J.L., 1994. Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature*, **368**(6467), pp. 119-124.
- KONARSKA, M.M., GRABOWSKI, P.J., PADGETT, R.A. and SHARP, P.A., 1985. Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors. *Nature*, **313**(6003), pp. 552-557.
- KONARSKA, M.M., VILARDELL, J. and QUERY, C.C., 2006. Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Molecular cell*, **21**(4), pp. 543-553.
- KOSHKIN, A.A., SINGH, S.K., NIELSEN, P., RAJWANSHI, V.K., KUMAR, R., MELDGAARD, M., OLSEN, C.E., WENGEL, J., 1998. LNA (Locked Nucleic Acids): Synthesis of the Adenine, Cytosine, Guanine, 5-Methylcytosine, Thymine and Uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron*, **54**(14), pp. 3607-3630.
- KOTLAJICH, M.V., CRABB, T.L. and HERTEL, K.J., 2009. Spliceosome assembly pathways for different types of alternative splicing converge during commitment to splice site pairing in the A complex. *Molecular and cellular biology*, **29**(4), pp. 1072-1082.
- KRAINER, A.R., CONWAY, G.C. and KOZAK, D., 1990. The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. *Cell*, **62**(1), pp. 35-42.
- KRAMER, A., KELLER, W., APPEL, B. and LUHRMANN, R., 1984. The 5' terminus of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA precursors. *Cell*, **38**(1), pp. 299-307.

KRAMER, A. and UTANS, U., 1991. Three protein factors (SF1, SF3 and U2AF) function in pre-splicing complex formation in addition to snRNPs. *The EMBO journal*, **10**(6), pp. 1503-1509.

KUHN, A.N., VAN SANTEN, M.A., SCHWIENHORST, A., URLAUB, H. and LUHRMANN, R., 2009. Stalling of spliceosome assembly at distinct stages by small-molecule inhibitors of protein acetylation and deacetylation. *RNA (New York, N.Y.)*, **15**(1), pp. 153-175.

KYRIAKOPOULOU, C., LARSSON, P., LIU, L., SCHUSTER, J., SODERBOM, F., KIRSEBOM, L.A. and VIRTANEN, A., 2006. U1-like snRNAs lacking complementarity to canonical 5' splice sites. *RNA (New York, N.Y.)*, **12**(9), pp. 1603-1611.

LABOURIER, E., ADAMS, M.D. and RIO, D.C., 2001. Modulation of P-element pre-mRNA splicing by a direct interaction between PSI and U1 snRNP 70K protein. *Molecular cell*, **8**(2), pp. 363-373.

LAKOWICZ, J.R., 1999. Principles of fluorescence spectroscopy 2nd edition, New York, Kluwer Academic/Plenum Publishers.

LALLENA, M.J., CHALMERS, K.J., LLAMAZARES, S., LAMOND, A.I. and VALCARCEL, J., 2002. Splicing regulation at the second catalytic step by Sex-lethal involves 3' splice site recognition by SPF45. *Cell*, **109**(3), pp. 285-296.

LAM, B.J. and HERTEL, K.J., 2002. A general role for splicing enhancers in exon definition. *RNA (New York, N.Y.)*, **8**(10), pp. 1233-1241.

LANDER, E.S., LINTON, L.M., BIRREN, B., NUSBAUM, C., ZODY, M.C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K., MELDRIM, J., MESIROV, J.P., MIRANDA, C., MORRIS, W., NAYLOR, J., RAYMOND, C., ROSETTI, M., SANTOS, R., SHERIDAN, A., SOUGNEZ, C., STANGE-THOMANN, N., STOJANOVIC, N., SUBRAMANIAN, A., WYMAN, D., ROGERS, J., SULSTON, J., AINSCOUGH, R., BECK, S., BENTLEY, D., BURTON, J., CLEE, C., CARTER, N., COULSON, A., DEADMAN, R., DELOUKAS, P., DUNHAM, A., DUNHAM, I., DURBIN, R., FRENCH, L., GRAFHAM, D., GREGORY, S., HUBBARD, T., HUMPHRAY, S., HUNT, A., JONES, M., LLOYD, C., MCMURRAY, A., MATTHEWS, L., MERCER, S., MILNE, S., MULLIKIN, J.C., MUNGALL, A., PLUMB, R., ROSS, M., SHOWNKEEN, R., SIMS, S., WATERSTON, R.H., WILSON, R.K., HILLIER, L.W., MCPHERSON, J.D., MARRA, M.A., MARDIS, E.R., FULTON, L.A., CHINWALLA, A.T., PEPIN, K.H., GISH, W.R., CHISSOE, S.L., WENDL, M.C., DELEHAUNTY, K.D., MINER, T.L., DELEHAUNTY, A., KRAMER, J.B., COOK, L.L., FULTON, R.S., JOHNSON, D.L., MINX, P.J., CLIFTON, S.W., HAWKINS, T., BRANSCOMB, E., PREDKI, P., RICHARDSON, P., WENNING, S., SLEZAK, T., DOGGETT, N., CHENG, J.F., OLSEN, A., LUCAS, S., ELKIN, C., UBERBACHER, E., FRAZIER, M., GIBBS, R.A., MUZNY, D.M., SCHERER, S.E., BOUCK, J.B., SODERGREN, E.J., WORLEY, K.C., RIVES, C.M., GORRELL, J.H., METZKER, M.L., NAYLOR, S.L., KUCHERLAPATI,

R.S., NELSON, D.L., WEINSTOCK, G.M., SAKAKI, Y., FUJIYAMA, A., HATTORI, M., YADA, T., TOYODA, A., ITOH, T., KAWAGOE, C., WATANABE, H., TOTOKI, Y., TAYLOR, T., WEISSENBACH, J., HEILIG, R., SAURIN, W., ARTIGUENAVE, F., BROTTIER, P., BRULS, T., PELLETIER, E., ROBERT, C., WINCKER, P., SMITH, D.R., DOUCETTE-STAMM, L., RUBENFIELD, M., WEINSTOCK, K., LEE, H.M., DUBOIS, J., ROSENTHAL, A., PLATZER, M., NYAKATURA, G., TAUDIEN, S., RUMP, A., YANG, H., YU, J., WANG, J., HUANG, G., GU, J., HOOD, L., ROWEN, L., MADAN, A., QIN, S., DAVIS, R.W., FEDERSPIEL, N.A., ABOLA, A.P., PROCTOR, M.J., MYERS, R.M., SCHMUTZ, J., DICKSON, M., GRIMWOOD, J., COX, D.R., OLSON, M.V., KAUL, R., RAYMOND, C., SHIMIZU, N., KAWASAKI, K., MINOSHIMA, S., EVANS, G.A., ATHANASIOU, M., SCHULTZ, R., ROE, B.A., CHEN, F., PAN, H., RAMSER, J., LEHRACH, H., REINHARDT, R., MCCOMBIE, W.R., DE LA BASTIDE, M., DEDHIA, N., BLOCKER, H., HORNISCHER, K., NORDSIEK, G., AGARWALA, R., ARAVIND, L., BAILEY, J.A., BATEMAN, A., BATZOGLOU, S., BIRNEY, E., BORK, P., BROWN, D.G., BURGE, C.B., CERUTTI, L., CHEN, H.C., CHURCH, D., CLAMP, M., COPLEY, R.R., DOERKS, T., EDDY, S.R., EICHLER, E.E., FUREY, T.S., GALAGAN, J., GILBERT, J.G., HARMON, C., HAYASHIZAKI, Y., HAUSSLER, D., HERMJAKOB, H., HOKAMP, K., JANG, W., JOHNSON, L.S., JONES, T.A., KASIF, S., KASPRYZK, A., KENNEDY, S., KENT, W.J., KITTS, P., KOONIN, E.V., KORF, I., KULP, D., LANCET, D., LOWE, T.M., MCCLYSAGHT, A., MIKKELSEN, T., MORAN, J.V., MULDER, N., POLLARA, V.J., PONTING, C.P., SCHULER, G., SCHULTZ, J., SLATER, G., SMIT, A.F., STUPKA, E., SZUSTAKOWSKI, J., THIERRY-MIEG, D., THIERRY-MIEG, J., WAGNER, L., WALLIS, J., WHEELER, R., WILLIAMS, A., WOLF, Y.I., WOLFE, K.H., YANG, S.P., YEH, R.F., COLLINS, F., GUYER, M.S., PETERSON, J., FELSENFELD, A., WETTERSTRAND, K.A., PATRINOS, A., MORGAN, M.J., DE JONG, P., CATANESE, J.J., OSOEGAWA, K., SHIZUYA, H., CHOI, S., CHEN, Y.J. and INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM, 2001. Initial sequencing and analysis of the human genome. *Nature*, **409**(6822), pp. 860-921.

LAREAU, L.F., INADA, M., GREEN, R.E., WENGROD, J.C. and BRENNER, S.E., 2007. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature*, **446**(7138), pp. 926-929.

LEAKE, M.C., CHANDLER, J.H., WADHAMS, G.H., BAI, F., BERRY, R.M. and ARMITAGE, J.P., 2006. Stoichiometry and turnover in single, functioning membrane protein complexes. *Nature*, **443**(7109), pp. 355-358.

LEAR, A.L., EPERON, L.P., WHEATLEY, I.M. and EPERON, I.C., 1990. Hierarchy for 5' splice site preference determined in vivo. *Journal of Molecular Biology*, **211**(1), pp. 103-115.

LEE, K.A., BINDEREIF, A. and GREEN, M.R., 1988. A small-scale procedure for preparation of nuclear extracts that support efficient transcription and pre-mRNA splicing. *Gene analysis techniques*, **5**(2), pp. 22-31.

LEGRAIN, P. and ROSBASH, M., 1989. Some cis- and trans-acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell*, **57**(4), pp. 573-583.

- LERNER, M.R., BOYLE, J.A., MOUNT, S.M., WOLIN, S.L. and STEITZ, J.A., 1980. Are snRNPs involved in splicing? *Nature*, **283**(5743), pp. 220-224.
- LEV-MAOR, G., SOREK, R., SHOMRON, N. and AST, G., 2003. The birth of an alternatively spliced exon: 3' splice-site selection in Alu exons. *Science (New York, N.Y.)*, **300**(5623), pp. 1288-1291.
- LEWIS, J.D., IZAURRALDE, E., JARMOLOWSKI, A., MCGUIGAN, C. and MATTAJ, I.W., 1996. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes & development*, **10**(13), pp. 1683-1698.
- LI, Y. and BLENCOWE, B.J., 1999. Distinct factor requirements for exonic splicing enhancer function and binding of U2AF to the polypyrimidine tract. *The Journal of biological chemistry*, **274**(49), pp. 35074-35079.
- LIM, S.R. and HERTEL, K.J., 2004. Commitment to splice site pairing coincides with A complex formation. *Molecular cell*, **15**(3), pp. 477-483.
- LIU, Z., LUYTEN, I., BOTTOMLEY, M.J., MESSIAS, A.C., HOUNGNINOU-MOLANGO, S., SPRANGERS, R., ZANIER, K., KRAMER, A. and SATTler, M., 2001. Structural basis for recognition of the intron branch site RNA by splicing factor 1. *Science (New York, N.Y.)*, **294**(5544), pp. 1098-1102.
- LIU, Z.R., 2002. p68 RNA helicase is an essential human splicing factor that acts at the U1 snRNA-5' splice site duplex. *Molecular and cellular biology*, **22**(15), pp. 5443-5450.
- LOEB, D.D., MACK, A.A. and TIAN, R., 2002. A secondary structure that contains the 5' and 3' splice sites suppresses splicing of duck hepatitis B virus pregenomic RNA. *Journal of virology*, **76**(20), pp. 10195-10202.
- LONG, J.C. and CACERES, J.F., 2009. The SR protein family of splicing factors: master regulators of gene expression. *The Biochemical journal*, **417**(1), pp. 15-27.
- LONGMAN, D., JOHNSTONE, I.L. and CACERES, J.F., 2000. Functional characterization of SR and SR-related genes in *Caenorhabditis elegans*. *The EMBO journal*, **19**(7), pp. 1625-1637.
- LOOMIS, R.J., NAOE, Y., PARKER, J.B., SAVIC, V., BOZOVSKY, M.R., MACFARLAN, T., MANLEY, J.L. and CHAKRAVARTI, D., 2009. Chromatin binding of SRp20 and ASF/SF2 and dissociation from mitotic chromosomes is modulated by histone H3 serine 10 phosphorylation. *Molecular cell*, **33**(4), pp. 450-461.
- LOPEZ-BIGAS, N., AUDIT, B., OUZOUNIS, C., PARRA, G. and GUIGO, R., 2005. Are splicing mutations the most frequent cause of hereditary disease? *FEBS letters*, **579**(9), pp. 1900-1903.

- LUND, E. and DAHLBERG, J.E., 1987. Differential accumulation of U1 and U4 small nuclear RNAs during *Xenopus* development. *Genes & development*, **1**(1), pp. 39-46.
- LUND, M. and KJEMS, J., 2002. Defining a 5' splice site by functional selection in the presence and absence of U1 snRNA 5' end. *RNA (New York, N.Y.)*, **8**(2), pp. 166-179.
- MAKAROV, E.M., MAKAROVA, O.V., URLAUB, H., GENTZEL, M., WILL, C.L., WILM, M. and LUHRMANN, R., 2002. Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science (New York, N.Y.)*, **298**(5601), pp. 2205-2208.
- MATHEW, R., HARTMUTH, K., MOHLMANN, S., URLAUB, H., FICNER, R. and LUHRMANN, R., 2008. Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. *Nature structural & molecular biology*, **15**(5), pp. 435-443.
- MATZ, M.V., FRADKOV, A.F., LABAS, Y.A., SAVITSKY, A.P., ZARAIISKY, A.G., MARKELOV, M.L. and LUKYANOV, S.A., 1999. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature biotechnology*, **17**(10), pp. 969-973.
- MARDON, H.J., SEBASTIO, G. and BARALLE, F.E., 1987. A role for exon sequences in alternative splicing of the human fibronectin gene. *Nucleic acids research*, **15**(19), pp. 7725-7733.
- MARTINEZ-CONTRERAS, R., CLOUTIER, P., SHKRETA, L., FISETTE, J.F., REVIL, T. and CHABOT, B., 2007. hnRNP proteins and splicing control. *Advances in Experimental Medicine and Biology*, **623**, pp. 123-147.
- MARTINEZ-CONTRERAS, R., FISETTE, J.F., NASIM, F.U., MADDEN, R., CORDEAU, M. and CHABOT, B., 2006. Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. *PLoS biology*, **4**(2), pp. e21.
- MARTINS DE ARAUJO, M., BONNAL, S., HASTINGS, M.L., KRAINER, A.R. and VALCARCEL, J., 2009. Differential 3' splice site recognition of SMN1 and SMN2 transcripts by U2AF and U2 snRNP. *RNA (New York, N.Y.)*, **15**(4), pp. 515-523.
- MATLIN, A.J., CLARK, F. and SMITH, C.W., 2005. Understanding alternative splicing: towards a cellular code. *Nature reviews. Molecular cell biology*, **6**(5), pp. 386-398.
- MAYEDA, A., HELFMAN, D.M. and KRAINER, A.R., 1993. Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. *Molecular and cellular biology*, **13**(5), pp. 2993-3001.
- MAYEDA, A. and KRAINER, A.R., 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell*, **68**(2), pp. 365-375.

MAYEDA, A., MUNROE, S.H., CACERES, J.F. and KRAINER, A.R., 1994. Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. *The EMBO journal*, **13**(22), pp. 5483-5495.

MAYEDA, A. and OHSHIMA, Y., 1988. Short donor site sequences inserted within the intron of beta-globin pre-mRNA serve for splicing in vitro. *Molecular and cellular biology*, **8**(10), pp. 4484-4491.

MCANANEY, T.B., ZENG, W., DOE, C.F., BHANJI, N., WAKELIN, S., PEARSON, D.S., ABBYAD, P., SHI, X., BOXER, S.G. and BAGSHAW, C.R., 2005. Protonation, photobleaching, and photoactivation of yellow fluorescent protein (YFP 10C): a unifying mechanism. *Biochemistry*, **44**(14), pp. 5510-5524.

MCLEAN, M.A., RAJFUR, Z., CHEN, Z., HUMPHREY, D., YANG, B., SLIGAR, S.G. and JACOBSON, K., 2009. Mechanism of chromophore assisted laser inactivation employing fluorescent proteins. *Analytical Chemistry*, **81**(5), pp. 1755-1761.

MERENDINO, L., GUTH, S., BILBAO, D., MARTINEZ, C. and VALCARCEL, J., 1999. Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature*, **402**(6763), pp. 838-841.

MERMOUD, J.E., COHEN, P. and LAMOND, A.I., 1992. Ser/Thr-specific protein phosphatases are required for both catalytic steps of pre-mRNA splicing. *Nucleic acids research*, **20**(20), pp. 5263-5269.

MERMOUD, J.E., COHEN, P.T. and LAMOND, A.I., 1994. Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. *The EMBO journal*, **13**(23), pp. 5679-5688.

MICHAUD, S. and REED, R., 1993. A functional association between the 5' and 3' splice site is established in the earliest prespliceosome complex (E) in mammals. *Genes & development*, **7**(6), pp. 1008-1020.

MICHAUD, S. and REED, R., 1991. An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes & development*, **5**(12B), pp. 2534-2546.

MISTELI, T., CACERES, J.F. and SPECTOR, D.L., 1997. The dynamics of a pre-mRNA splicing factor in living cells. *Nature*, **387**(6632), pp. 523-527.

MOORE, M.J. and SHARP, P.A., 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature*, **365**(6444), pp. 364-368.

- MOUNT, S.M., 1982. A catalogue of splice junction sequences. *Nucleic acids research*, **10**(2), pp. 459-472.
- MOUNT, S.M., PETTERSSON, I., HINTERBERGER, M., KARMAS, A. and STEITZ, J.A., 1983. The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. *Cell*, **33**(2), pp. 509-518.
- MOUNT, S.M. and STEITZ, J.A., 1981. Sequence of U1 RNA from *Drosophila melanogaster*: implications for U1 secondary structure and possible involvement in splicing. *Nucleic acids research*, **9**(23), pp. 6351-6368.
- MULLEN, M.P., SMITH, C.W., PATTON, J.G. and NADAL-GINARD, B., 1991. Alpha-tropomyosin mutually exclusive exon selection: competition between branchpoint/polypyrimidine tracts determines default exon choice. *Genes & development*, **5**(4), pp. 642-655.
- MULLER, S., WOLPENSINGER, B., ANGENITZKI, M., ENGEL, A., SPERLING, J. and SPERLING, R., 1998. A supraspliceosome model for large nuclear ribonucleoprotein particles based on mass determinations by scanning transmission electron microscopy. *Journal of Molecular Biology*, **283**(2), pp. 383-394.
- NASIM, F.U., HUTCHISON, S., CORDEAU, M. and CHABOT, B., 2002. High-affinity hnRNP A1 binding sites and duplex-forming inverted repeats have similar effects on 5' splice site selection in support of a common looping out and repression mechanism. *RNA (New York, N.Y.)*, **8**(8), pp. 1078-1089.
- NELISSEN, R.L., WILL, C.L., VAN VENROOIJ, W.J. and LUHRMANN, R., 1994. The association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in part by common U snRNP proteins. *The EMBO journal*, **13**(17), pp. 4113-4125.
- NELSON, K.K. and GREEN, M.R., 1990. Mechanism for cryptic splice site activation during pre-mRNA splicing. *Proceedings of the National Academy of Sciences of the United States of America*, **87**(16), pp. 6253-6257.
- NELSON, K.K. and GREEN, M.R., 1988. Splice site selection and ribonucleoprotein complex assembly during in vitro pre-mRNA splicing. *Genes & development*, **2**(3), pp. 319-329.
- NEUENKIRCHEN, N., CHARI, A. and FISCHER, U., 2008. Deciphering the assembly pathway of Sm-class U snRNPs. *FEBS letters*, **582**(14), pp. 1997-2003.
- NI, J.Z., GRATE, L., DONOHUE, J.P., PRESTON, C., NOBIDA, N., O'BRIEN, G., SHIUE, L., CLARK, T.A., BLUME, J.E. and ARES, M., JR, 2007. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes & development*, **21**(6), pp. 708-718.

- NIFOSI, R., FERRARI, A., ARCANGELI, C., TOZZINI, V., PELLEGRINI, V., BELTRAM, F., 2003. Photoreversible dark state in a tristable green fluorescent protein variant. *Journal of Physical Chemistry B*, **107**, pp. 1679-1684.
- NIWA, H., INOUE, S., HIRANO, T., MATSUNO, T., KOJIMA, S., KUBOTA, M., OHASHI, M. and TSUJI, F.I., 1996. Chemical nature of the light emitter of the Aequorea green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, **93**(24), pp. 13617-13622.
- NOBLE, J.C., PAN, Z.Q., PRIVES, C. and MANLEY, J.L., 1987. Splicing of SV40 early pre-mRNA to large T and small t mRNAs utilizes different patterns of lariat branch sites. *Cell*, **50**(2), pp. 227-236.
- OBERSTRASS, F.C., AUWETER, S.D., ERAT, M., HARGOUS, Y., HENNING, A., WENTER, P., REYMOND, L., AMIR-AHMADY, B., PITSCH, S., BLACK, D.L. and ALLAIN, F.H., 2005. Structure of PTB bound to RNA: specific binding and implications for splicing regulation. *Science (New York, N.Y.)*, **309**(5743), pp. 2054-2057.
- OHE, K. and MAYEDA, A., 2010. HMGA1a trapping of U1 snRNP at an authentic 5' splice site induces aberrant exon skipping in sporadic Alzheimer's disease. *Molecular and cellular biology*, **30**(9), pp. 2220-2228.
- OKUNOLA, H.L. and KRAINER, A.R., 2009. Cooperative-binding and splicing-repressive properties of hnRNP A1. *Molecular and cellular biology*, **29**(20), pp. 5620-5631.
- O'MULLANE, L. and EPERON, I.C., 1998. The pre-mRNA 5' cap determines whether U6 small nuclear RNA succeeds U1 small nuclear ribonucleoprotein particle at 5' splice sites. *Molecular and cellular biology*, **18**(12), pp. 7510-7520.
- OUBRIDGE, C., ITO, N., EVANS, P.R., TEO, C.H. and NAGAI, K., 1994. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**(6505), pp. 432-438.
- PADGETT, R.A., MOUNT, S.M., STEITZ, J.A. and SHARP, P.A., 1983. Splicing of messenger RNA precursors is inhibited by antisera to small nuclear ribonucleoprotein. *Cell*, **35**(1), pp. 101-107.
- PARK, J.W., PARISKY, K., CELOTTO, A.M., REENAN, R.A. and GRAVELEY, B.R., 2004. Identification of alternative splicing regulators by RNA interference in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(45), pp. 15974-15979.

- PARRA, M.K., TAN, J.S., MOHANDAS, N. and CONBOY, J.G., 2008. Intrasplicing coordinates alternative first exons with alternative splicing in the protein 4.1R gene. *The EMBO journal*, **27**(1), pp. 122-131.
- PASMAN, Z. and GARCIA-BLANCO, M.A., 1996. The 5' and 3' splice sites come together via a three dimensional diffusion mechanism. *Nucleic acids research*, **24**(9), pp. 1638-1645.
- PATEL, S.B. and BELLINI, M., 2008. The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic acids research*, **36**(20), pp. 6482-6493.
- PATTON, J.R. and PEDERSON, T., 1988. The Mr 70,000 protein of the U1 small nuclear ribonucleoprotein particle binds to the 5' stem-loop of U1 RNA and interacts with Sm domain proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **85**(3), pp. 747-751.
- PEREZ, I., LIN, C.H., MCAFEE, J.G. and PATTON, J.G., 1997. Mutation of PTB binding sites causes misregulation of alternative 3' splice site selection in vivo. *RNA (New York, N.Y.)*, **3**(7), pp. 764-778.
- POLLARD, A.J., KRAINER, A.R., ROBSON, S.C. and EUROPE-FINNER, G.N., 2002. Alternative splicing of the adenylyl cyclase stimulatory G-protein G alpha(s) is regulated by SF2/ASF and heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) and involves the use of an unusual TG 3'-splice Site. *The Journal of biological chemistry*, **277**(18), pp. 15241-15251.
- POLLARD, A.J., SPAREY, C., ROBSON, S.C., KRAINER, A.R. and EUROPE-FINNER, G.N., 2000. Spatio-temporal expression of the trans-acting splicing factors SF2/ASF and heterogeneous ribonuclear proteins A1/A1B in the myometrium of the pregnant human uterus: a molecular mechanism for regulating regional protein isoform expression in vivo. *The Journal of clinical endocrinology and metabolism*, **85**(5), pp. 1928-1936.
- POMERANZ-KRUMMEL, D.A., OUBRIDGE, C., LEUNG, A.K., LI, J. and NAGAI, K., 2009. Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Nature*, **458**(7237), pp. 475-480.
- QUERY, C.C., MOORE, M.J. and SHARP, P.A., 1994. Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. *Genes & development*, **8**(5), pp. 587-597.
- RAKER, V.A., HARTMUTH, K., KASTNER, B. and LUHRMANN, R., 1999. Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. *Molecular and cellular biology*, **19**(10), pp. 6554-6565.

- REED, R., 2000. Mechanisms of fidelity in pre-mRNA splicing. *Current opinion in cell biology*, **12**(3), pp. 340-345.
- REED, R., 1990. Protein composition of mammalian spliceosomes assembled in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, **87**(20), pp. 8031-8035.
- REED, R., 1989. The organization of 3' splice-site sequences in mammalian introns. *Genes & development*, **3**(12B), pp. 2113-2123.
- REED, R. and MANIATIS, T., 1986. A role for exon sequences and splice-site proximity in splice-site selection. *Cell*, **46**(5), pp. 681-690.
- REED, R. and MANIATIS, T., 1985. Intron sequences involved in lariat formation during pre-mRNA splicing. *Cell*, **41**(1), pp. 95-105.
- REICHERT, V. and MOORE, M.J., 2000. Better conditions for mammalian in vitro splicing provided by acetate and glutamate as potassium counterions. *Nucleic acids research*, **28**(2), pp. 416-423.
- RIPPE, K., 2001. Making contacts on a nucleic acid polymer. *Trends in biochemical sciences*, **26**(12), pp. 733-740.
- ROBBERSON, B.L., COTE, G.J. and BERGET, S.M., 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Molecular and cellular biology*, **10**(1), pp. 84-94.
- ROCA, X. and KRAINER, A.R., 2009. Recognition of atypical 5' splice sites by shifted base-pairing to U1 snRNA. *Nature structural & molecular biology*, **16**(2), pp. 176-182.
- ROCA, X., SACHIDANANDAM, R. and KRAINER, A.R., 2005. Determinants of the inherent strength of human 5' splice sites. *RNA (New York, N.Y.)*, **11**(5), pp. 683-698.
- ROGERS, J. and WALL, R., 1980. A mechanism for RNA splicing. *Proceedings of the National Academy of Sciences of the United States of America*, **77**(4), pp. 1877-1879.
- ROMANO, M., MARCUCCI, R., BURATTI, E., AYALA, Y.M., SEBASTIO, G. and BARALLE, F.E., 2002. Regulation of 3' splice site selection in the 844ins68 polymorphism of the cystathionine Beta -synthase gene. *The Journal of biological chemistry*, **277**(46), pp. 43821-43829.
- ROOKE, N., MARKOVTSOV, V., CAGAVI, E. and BLACK, D.L., 2003. Roles for SR proteins and hnRNP A1 in the regulation of c-src exon N1. *Molecular and cellular biology*, **23**(6), pp. 1874-1884.

- ROSCIGNO, R.F. and GARCIA-BLANCO, M.A., 1995. SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA (New York, N.Y.)*, **1**(7), pp. 692-706.
- ROSCIGNO, R.F., WEINER, M. and GARCIA-BLANCO, M.A., 1993. A mutational analysis of the polypyrimidine tract of introns. Effects of sequence differences in pyrimidine tracts on splicing. *The Journal of biological chemistry*, **268**(15), pp. 11222-11229.
- ROSEL, T.D., HUNG, L.H., MEDENBACH, J., DONDE, K., STARKE, S., BENES, V., RATSCH, G. and BINDEREIF, A., 2011. RNA-Seq analysis in mutant zebrafish reveals role of U1C protein in alternative splicing regulation. *The EMBO journal*, **30**(10), pp. 1965-1976.
- ROSSI, F., FORNE, T., ANTOINE, E., TAZI, J., BRUNEL, C. and CATHALA, G., 1996. Involvement of U1 small nuclear ribonucleoproteins (snRNP) in 5' splice site-U1 snRNP interaction. *The Journal of biological chemistry*, **271**(39), pp. 23985-23991.
- RUBY, S.W., CHANG, T.H. and ABELSON, J., 1993. Four yeast spliceosomal proteins (PRP5, PRP9, PRP11, and PRP21) interact to promote U2 snRNP binding to pre-mRNA. *Genes & development*, **7**(10), pp. 1909-1925.
- RUDNER, D.Z., KANAAR, R., BREGER, K.S. and RIO, D.C., 1998. Interaction between subunits of heterodimeric splicing factor U2AF is essential in vivo. *Molecular and cellular biology*, **18**(4), pp. 1765-1773.
- RUSKIN, B., GREENE, J.M. and GREEN, M.R., 1985. Cryptic branch point activation allows accurate in vitro splicing of human beta-globin intron mutants. *Cell*, **41**(3), pp. 833-844.
- RUSKIN, B., KRAINER, A.R., MANIATIS, T. and GREEN, M.R., 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell*, **38**(1), pp. 317-331.
- RUSKIN, B., ZAMORE, P.D. and GREEN, M.R., 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell*, **52**(2), pp. 207-219.
- SALTZMAN, A.L., PAN, Q. and BLENCOWE, B.J., 2011. Regulation of alternative splicing by the core spliceosomal machinery. *Genes & development*, **25**(4), pp. 373-384.
- SAULIERE, J., SUREAU, A., EXPERT-BEZANCON, A. and MARIE, J., 2006. The polypyrimidine tract binding protein (PTB) represses splicing of exon 6B from the beta-tropomyosin pre-mRNA by directly interfering with the binding of the U2AF65 subunit. *Molecular and cellular biology*, **26**(23), pp. 8755-8769.
- SCHERLY, D., BOELEN, W., VAN VENROOIJ, W.J., DATHAN, N.A., HAMM, J. and MATTAJ, I.W., 1989. Identification of the RNA binding segment of human U1 A protein and definition of its binding site on U1 snRNA. *The EMBO journal*, **8**(13), pp. 4163-4170.

- SCHNEIDER, M., WILL, C.L., ANOKHINA, M., TAZI, J., URLAUB, H. and LUHRMANN, R., 2010. Exon definition complexes contain the tri-snRNP and can be directly converted into B-like precatalytic splicing complexes. *Molecular cell*, **38**(2), pp. 223-235.
- SCREATON, G.R., CACERES, J.F., MAYEDA, A., BELL, M.V., PLEBANSKI, M., JACKSON, D.G., BELL, J.I. and KRAINER, A.R., 1995. Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *The EMBO journal*, **14**(17), pp. 4336-4349.
- SELENKO, P., GREGOROVIC, G., SPRANGERS, R., STIER, G., RHANI, Z., KRAMER, A. and SATTLER, M., 2003. Structural basis for the molecular recognition between human splicing factors U2AF65 and SF1/mBBP. *Molecular cell*, **11**(4), pp. 965-976.
- SELVAKUMAR, M. and HELFMAN, D.M., 1999. Exonic splicing enhancers contribute to the use of both 3' and 5' splice site usage of rat beta-tropomyosin pre-mRNA. *RNA (New York, N.Y.)*, **5**(3), pp. 378-394.
- SHANER, N.C., STEINBACH, P.A. and TSIEN, R.Y., 2005. A guide to choosing fluorescent proteins. *Nature methods*, **2**(12), pp. 905-909.
- SHAPIRO, M.B. and SENAPATHY, P., 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic acids research*, **15**(17), pp. 7155-7174.
- SHARMA, S., KOHLSTAEDT, L.A., DAMIANOV, A., RIO, D.C. and BLACK, D.L., 2008. Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome. *Nature structural & molecular biology*, **15**(2), pp. 183-191.
- SHARMA, S., MARIS, C., ALLAIN, F.H. and BLACK, D.L., 2011. U1 snRNA directly interacts with polypyrimidine tract-binding protein during splicing repression. *Molecular cell*, **41**(5), pp. 579-588.
- SHEN, H. and GREEN, M.R., 2007. RS domain-splicing signal interactions in splicing of U12-type and U2-type introns. *Nature structural & molecular biology*, **14**(7), pp. 597-603.
- SHEN, H. and GREEN, M.R., 2006. RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. *Genes & development*, **20**(13), pp. 1755-1765.
- SHEN, H. and GREEN, M.R., 2004. A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. *Molecular cell*, **16**(3), pp. 363-373.

SHEN, H., KAN, J.L. and GREEN, M.R., 2004. Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. *Molecular cell*, **13**(3), pp. 367-376.

SHEN, J., ZU, K., CASS, C.L., BEYER, A.L. and HIRSH, J., 1995. Exon skipping by overexpression of a *Drosophila* heterogeneous nuclear ribonucleoprotein in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, **92**(6), pp. 1822-1825.

SHI, Y., REDDY, B. and MANLEY, J.L., 2006. PP1/PP2A phosphatases are required for the second step of Pre-mRNA splicing and target specific snRNP proteins. *Molecular cell*, **23**(6), pp. 819-829.

SHU, D., ZHANG, H., JIN, J. and GUO, P., 2007. Counting of six pRNAs of phi29 DNA-packaging motor with customized single-molecule dual-view system. *The EMBO journal*, **26**(2), pp. 527-537.

SHUKLA, S., DEL GATTO-KONCZAK, F., BREATHNACH, R. and FISHER, S.A., 2005. Competition of PTB with TIA proteins for binding to a U-rich cis-element determines tissue-specific splicing of the myosin phosphatase targeting subunit 1. *RNA (New York, N.Y.)*, **11**(11), pp. 1725-1736.

SICKMIER, E.A., FRATO, K.E., SHEN, H., PARANAWITHANA, S.R., GREEN, M.R. and KIELKOPF, C.L., 2006. Structural basis for polypyrimidine tract recognition by the essential pre-mRNA splicing factor U2AF65. *Molecular cell*, **23**(1), pp. 49-59.

SIMARD, M.J. and CHABOT, B., 2002. SRp30c is a repressor of 3' splice site utilization. *Molecular and cellular biology*, **22**(12), pp. 4001-4010.

SINGH, N.N., SINGH, R.N. and ANDROPHY, E.J., 2007. Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes. *Nucleic acids research*, **35**(2), pp. 371-389.

SINGH, R., VALCARCEL, J. and GREEN, M.R., 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science (New York, N.Y.)*, **268**(5214), pp. 1173-1176.

SKORDIS, L.A., DUNCKLEY, M.G., YUE, B., EPERON, I.C. and MUNTONI, F., 2003. Bifunctional antisense oligonucleotides provide a trans-acting splicing enhancer that stimulates SMN2 gene expression in patient fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America*, **100**(7), pp. 4114-4119.

SMITH, C.W., CHU, T.T. and 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), pp. 4939-4952.

SMITH, C.W. and NADAL-GINARD, B., 1989. Mutually exclusive splicing of alpha-tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. *Cell*, **56**(5), pp. 749-758.

SMITH, C.W., PORRO, E.B., PATTON, J.G. and NADAL-GINARD, B., 1989. Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. *Nature*, **342**(6247), pp. 243-247.

SMITH, C.W. and VALCARCEL, J., 2000. Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends in biochemical sciences*, **25**(8), pp. 381-388.

SMITH, D.J., QUERY, C.C. and KONARSKA, M.M., 2008. "Nought may endure but mutability": spliceosome dynamics and the regulation of splicing. *Molecular cell*, **30**(6), pp. 657-666.

SOLNICK, D. and LEE, S.I., 1987. Amount of RNA secondary structure required to induce an alternative splice. *Molecular and cellular biology*, **7**(9), pp. 3194-3198.

SOREK, R., LEV-MAOR, G., REZNIK, M., DAGAN, T., BELINKY, F., GRAUR, D. and AST, G., 2004. Minimal conditions for exonization of intronic sequences: 5' splice site formation in alu exons. *Molecular cell*, **14**(2), pp. 221-231.

SOSNOWSKI, B.A., BELOTE, J.M. and MCKEOWN, M., 1989. Sex-specific alternative splicing of RNA from the transformer gene results from sequence-dependent splice site blockage. *Cell*, **58**(3), pp. 449-459.

STAKNIS, D. and REED, R., 1994(a). Direct interactions between pre-mRNA and six U2 small nuclear ribonucleoproteins during spliceosome assembly. *Molecular and cellular biology*, **14**(5), pp. 2994-3005.

STAKNIS, D. and REED, R., 1994(b). SR proteins promote the first specific recognition of Pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. *Molecular and cellular biology*, **14**(11), pp. 7670-7682.

STALEY, J.P. and GUTHRIE, C., 1999. An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. *Molecular cell*, **3**(1), pp. 55-64.

STAMM, S., 2002. Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome. *Human molecular genetics*, **11**(20), pp. 2409-2416.

STARK, H., DUBE, P., LUHRMANN, R. and KASTNER, B., 2001. Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature*, **409**(6819), pp. 539-542.

STEVENS, S.W., BARTA, I., GE, H.Y., MOORE, R.E., YOUNG, M.K., LEE, T.D. and ABELSON, J., 2001. Biochemical and genetic analyses of the U5, U6, and U4/U6 x U5 small nuclear ribonucleoproteins from *Saccharomyces cerevisiae*. *RNA (New York, N.Y.)*, **7**(11), pp. 1543-1553.

SULLIVAN, K.F., KAY, S.A., 1999. *Methods in Cell Biology*, Volume 58, Green Fluorescent Proteins, London, Academic Press.

SUN, Q., MAYEDA, A., HAMPSON, R.K., KRAINER, A.R. and ROTTMAN, F.M., 1993. General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. *Genes & development*, **7**(12B), pp. 2598-2608.

SUN, S., ZHANG, Z., SINHA, R., KARNI, R. and KRAINER, A.R., 2010. SF2/ASF autoregulation involves multiple layers of post-transcriptional and translational control. *Nature structural & molecular biology*, **17**(3), pp. 306-312.

TALERICO, M. and BERGET, S.M., 1990. Effect of 5' splice site mutations on splicing of the preceding intron. *Molecular and cellular biology*, **10**(12), pp. 6299-6305.

TARN, W.Y. and STEITZ, J.A., 1994. SR proteins can compensate for the loss of U1 snRNP functions in vitro. *Genes & development*, **8**(22), pp. 2704-2717.

TATEI, K., KIMURA, K. and OHSHIMA, Y., 1989. New methods to investigate ATP requirement for pre-mRNA splicing: inhibition by hexokinase/glucose or an ATP-binding site blocker. *Journal of Biochemistry*, **106**(3), pp. 372-375.

TEIGELKAMP, S., MUNDT, C., ACHSEL, T., WILL, C.L. and LUHRMANN, R., 1997. The human U5 snRNP-specific 100-kD protein is an RS domain-containing, putative RNA helicase with significant homology to the yeast splicing factor Prp28p. *RNA (New York, N.Y.)*, **3**(11), pp. 1313-1326.

TIAN, M. and MANIATIS, T., 1994. A splicing enhancer exhibits both constitutive and regulated activities. *Genes & development*, **8**(14), pp. 1703-1712.

TIAN, M. and MANIATIS, T., 1993. A splicing enhancer complex controls alternative splicing of doublesex pre-mRNA. *Cell*, **74**(1), pp. 105-114.

- TIAN, M. and MANIATIS, T., 1992. Positive control of pre-mRNA splicing in vitro. *Science (New York, N.Y.)*, **256**(5054), pp. 237-240.
- TOOR, N., KEATING, K.S., TAYLOR, S.D. and PYLE, A.M., 2008. Crystal structure of a self-spliced group II intron. *Science (New York, N.Y.)*, **320**(5872), pp. 77-82.
- TREISMAN, R., ORKIN, S.H. and MANIATIS, T., 1983. Specific transcription and RNA splicing defects in five cloned beta-thalassaemia genes. *Nature*, **302**(5909), pp. 591-596.
- TSAI, K.W., TSENG, H.C. and LIN, W.C., 2008. Two wobble-splicing events affect ING4 protein subnuclear localization and degradation. *Experimental cell research*, **314**(17), pp. 3130-3141.
- ULBRICH, M.H. and ISACOFF, E.Y., 2007. Subunit counting in membrane-bound proteins. *Nature methods*, **4**(4), pp. 319-321.
- VALCARCEL, J., GAUR, R.K., SINGH, R. and GREEN, M.R., 1996. Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected. *Science (New York, N.Y.)*, **273**(5282), pp. 1706-1709.
- VALCARCEL, J., SINGH, R., ZAMORE, P.D. and GREEN, M.R., 1993. The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. *Nature*, **362**(6416), pp. 171-175.
- VALENCIA, P., DIAS, A.P. and REED, R., 2008. Splicing promotes rapid and efficient mRNA export in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(9), pp. 3386-3391.
- VARANI, L., GUNDERSON, S.I., MATTAJ, I.W., KAY, L.E., NEUHAUS, D. and VARANI, G., 2000. The NMR structure of the 38 kDa U1A protein - PIE RNA complex reveals the basis of cooperativity in regulation of polyadenylation by human U1A protein. *Nature structural biology*, **7**(4), pp. 329-335.
- VICKERS, T.A., SABRIPOUR, M. and CROOKE, S.T., 2011. U1 adaptors result in reduction of multiple pre-mRNA species principally by sequestering U1snRNP. *Nucleic acids research*, **39**(10), pp. e71.
- WAHL, M.C., WILL, C.L. and LUHRMANN, R., 2009. The spliceosome: design principles of a dynamic RNP machine. *Cell*, **136**(4), pp. 701-718.
- WANG, E.T., SANDBERG, R., LUO, S., KHREBTUKOVA, I., ZHANG, L., MAYR, C., KINGSMORE, S.F., SCHROTH, G.P. and BURGE, C.B., 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature*, **456**(7221), pp. 470-476.

- WANG, J., XIAO, S.H. and MANLEY, J.L., 1998. Genetic analysis of the SR protein ASF/SF2: interchangeability of RS domains and negative control of splicing. *Genes & development*, **12**(14), pp. 2222-2233.
- WANG, Z., HOFFMANN, H.M. and GRABOWSKI, P.J., 1995. Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA (New York, N.Y.)*, **1**(1), pp. 21-35.
- WANG, Z., KAYIKCI, M., BRIESE, M., ZARNACK, K., LUSCOMBE, N.M., ROT, G., ZUPAN, B., CURK, T. and ULE, J., 2010. iCLIP predicts the dual splicing effects of TIA-RNA interactions. *PLoS biology*, **8**(10), pp. e1000530.
- WARKOCKI, Z., ODENWALDER, P., SCHMITZOVA, J., PLATZMANN, F., STARK, H., URLAUB, H., FICNER, R., FABRIZIO, P. and LUHRMANN, R., 2009. Reconstitution of both steps of *Saccharomyces cerevisiae* splicing with purified spliceosomal components. *Nature structural & molecular biology*, **16**(12), pp. 1237-1243.
- WASSARMAN, D.A. and STEITZ, J.A., 1992. Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. *Science (New York, N.Y.)*, **257**(5078), pp. 1918-1925.
- WEBER, G., TROWITZSCH, S., KASTNER, B., LUHRMANN, R. and WAHL, M.C., 2010. Functional organization of the Sm core in the crystal structure of human U1 snRNP. *The EMBO journal*, **29**(24), pp. 4172-4184.
- WIECZOREK, D.F., SMITH, C.W. and NADAL-GINARD, B., 1988. The rat alpha-tropomyosin gene generates a minimum of six different mRNAs coding for striated, smooth, and nonmuscle isoforms by alternative splicing. *Molecular and cellular biology*, **8**(2), pp. 679-694.
- WILL, C.L. and LUHRMANN, R., 2010. Spliceosome Structure and Function. *Cold Spring Harbor perspectives in biology*, .
- WILL, C.L., RUMPLER, S., KLEIN GUNNEWIEK, J., VAN VENROOIJ, W.J. and LUHRMANN, R., 1996. In vitro reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E) spliceosomal complexes. *Nucleic acids research*, **24**(23), pp. 4614-4623.
- WU, J. and MANLEY, J.L., 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes & development*, **3**(10), pp. 1553-1561.
- WU, J.Y. and MANIATIS, T., 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*, **75**(6), pp. 1061-1070.

- WU, S., ROMFO, C.M., NILSEN, T.W. and GREEN, M.R., 1999. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature*, **402**(6763), pp. 832-835.
- XIAO, S.H. and MANLEY, J.L., 1997. Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes & development*, **11**(3), pp. 334-344.
- XIAO, X., WANG, Z., JANG, M. and BURGE, C.B., 2007. Coevolutionary networks of splicing cis-regulatory elements. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(47), pp. 18583-18588.
- XU, Q., MODREK, B. and LEE, C., 2002. Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. *Nucleic acids research*, **30**(17), pp. 3754-3766.
- XU, Y.Z., NEWNHAM, C.M., KAMEOKA, S., HUANG, T., KONARSKA, M.M. and QUERY, C.C., 2004. Prp5 bridges U1 and U2 snRNPs and enables stable U2 snRNP association with intron RNA. *The EMBO journal*, **23**(2), pp. 376-385.
- YANG, F., MOSS, L.G. and PHILLIPS, G.N., JR, 1996. The molecular structure of green fluorescent protein. *Nature biotechnology*, **14**(10), pp. 1246-1251.
- YANG, X., BANI, M.R., LU, S.J., ROWAN, S., BEN-DAVID, Y. and CHABOT, B., 1994. The A1 and A1B proteins of heterogeneous nuclear ribonucleoproteins modulate 5' splice site selection in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, **91**(15), pp. 6924-6928.
- YU, Y., MARONEY, P.A., DENKER, J.A., ZHANG, X.H., DYBKOV, O., LUHRMANN, R., JANKOWSKY, E., CHASIN, L.A. and NILSEN, T.W., 2008. Dynamic regulation of alternative splicing by silencers that modulate 5' splice site competition. *Cell*, **135**(7), pp. 1224-1236.
- YUO, C.Y. and WEINER, A.M., 1989. A U1 small nuclear ribonucleoprotein particle with altered specificity induces alternative splicing of an adenovirus E1A mRNA precursor. *Molecular and cellular biology*, **9**(8), pp. 3429-3437.
- ZACHARIAS, D.A., VIOLIN, J.D., NEWTON, A.C. and TSIEN, R.Y., 2002. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science (New York, N.Y.)*, **296**(5569), pp. 913-916.
- ZAHLER, A.M., DAMGAARD, C.K., KJEMS, J. and CAPUTI, M., 2004. SC35 and heterogeneous nuclear ribonucleoprotein A/B proteins bind to a juxtaposed exonic splicing enhancer/exonic splicing silencer element to regulate HIV-1 tat exon 2 splicing. *The Journal of biological chemistry*, **279**(11), pp. 10077-10084.

- ZAMORE, P.D. and GREEN, M.R., 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proceedings of the National Academy of Sciences of the United States of America*, **86**(23), pp. 9243-9247.
- ZAMORE, P.D., PATTON, J.G. and GREEN, M.R., 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature*, **355**(6361), pp. 609-614.
- ZHANG, W.J. and WU, J.Y., 1996. Functional properties of p54, a novel SR protein active in constitutive and alternative splicing. *Molecular and cellular biology*, **16**(10), pp. 5400-5408.
- ZHOU, A., OU, A.C., CHO, A., BENZ, E.J., JR and HUANG, S.C., 2008. Novel splicing factor RBM25 modulates Bcl-x pre-mRNA 5' splice site selection. *Molecular and cellular biology*, **28**(19), pp. 5924-5936.
- ZHU, J. and KRAINER, A.R., 2000. Pre-mRNA splicing in the absence of an SR protein RS domain. *Genes & development*, **14**(24), pp. 3166-3178.
- ZHU, J., MAYEDA, A. and KRAINER, A.R., 2001. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Molecular cell*, **8**(6), pp. 1351-1361.
- ZHUANG, Y., LEUNG, H. and WEINER, A.M., 1987. The natural 5' splice site of simian virus 40 large T antigen can be improved by increasing the base complementarity to U1 RNA. *Molecular and cellular biology*, **7**(8), pp. 3018-3020.
- ZHUANG, Y. and WEINER, A.M., 1986. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell*, **46**(6), pp. 827-835.
- ZHUANG, Y.A., GOLDSTEIN, A.M. and WEINER, A.M., 1989. UACUAAC is the preferred branch site for mammalian mRNA splicing. *Proceedings of the National Academy of Sciences of the United States of America*, **86**(8), pp. 2752-2756.
- ZORIO, D.A. and BLUMENTHAL, T., 1999. Both subunits of U2AF recognize the 3' splice site in *Caenorhabditis elegans*. *Nature*, **402**(6763), pp. 835-838.
- ZUO, P. and MANIATIS, T., 1996. The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. *Genes & development*, **10**(11), pp. 1356-1368.
- ZUO, P. and MANLEY, J.L., 1994. The human splicing factor ASF/SF2 can specifically recognize pre-mRNA 5' splice sites. *Proceedings of the National Academy of Sciences of the United States of America*, **91**(8), pp. 3363-3367.

