



UNIVERSITY OF  
**LEICESTER**

**Investigation of the Mechanisms of Action  
of Exonic Splicing Enhancers Using Single  
Molecule Methods**

by

**Alia Abdulaziz Alf**

**Thesis**

Submitted to the University of Leicester

for the degree of

**Doctor of Philosophy**

**Molecular and Cell Biology**

October 2016

# Abstract

**Investigation of the Mechanisms of Action of Exonic splicing Enhancers Using Single Molecule Methods - Alia Abdulaziz Alfi.**

The main purpose of studying the state of binding SRm160 at single molecule level is to extract fundamental information about this protein (SRm160) and the mechanism of its involvement in splicing reactions.

Single molecule methods have been widely used in colocalisation applications due to their advantages in following the state of a specific molecule (single molecule) of interest in real time. In the context of SRm160 protein, its involvement in splicing reaction and its relationship with other snRNPs (factors) remain important issues that may be involved in the early complex E and complex A, such that, it may induce a restructuring of the pre-mRNA complexes. Very few research papers were published regarding what is SRm160. Since then, it was believed that the binding of this protein to RNA is through other SR proteins as SRm160 does not consist of RRM motif domain. It was also believed that SRm160 binds as a bridge along side RNA. Regarding all studies on SRm160, we found it interesting to study the state/ the behaviour of this protein at single molecule level in order to elucidate its mechanism in real time. This thesis identified new insight for how SRm160 protein interact with RNA and how other snRNPs might affect this reaction. This identifica-

## ABSTRACT

---

tion is based on single molecule technique that depend on counting bleaching steps of the attached fluorescent fluorophores (colocalisation measurements). Additionally, the project was extended to initiate a FRET analysis in order to investigate whether the bound-SRm160 affected the flexibility movement of a substrate or not.

Two different approaches are pursued:

- a. The fifth chapter proposes a novel single molecule technique which is introduced as single molecule colocalization system. This system is using a suitable microscope for excitation via total internal reflection (TIRF) to obtain the single molecule data. Two different fluorophores (the chosen fluorophores are mEGFP and Cy5) were implemented to allow the investigation of the interaction between SRm160 molecules and the pre-mRNA during splicing reaction. Multiple pre-mRNAs were used, such as, GloC and SMN2 constructs. Multiple conditions were tested, such as, the absence of ATP, the presence of ATP and the presence of other factors (anti oligonucleotides) which might affect the complexes formation or even snRNPs binding. Following this, we can propose a new paradigm for how SRm160 interacts with pre-mRNA. The obtained results demonstrate that the association of GFP-SRm160 with SMN2 substrate is increased in the presence of the ESEs site. This association is not affected in the presence of anti-U1 oligonucleotide. The distribution of this association is significantly reduced with a single SRm160 protein remaining on the substrate in the presence of anti-U2 snRNP. These results demonstrate various points; the association of SRm160 with SMN2 pre-mRNA is enhanced via ESE sites, and is highly dependent on U2 snRNP but not U1 snRNP. These results refute the speculation in the literature that the association of SRm160 is highly dependent on U1

## ABSTRACT

---

snRNP and stabilized by U2 snRNP.

- b. The sixth chapter of the thesis proposes FRET system in order to allow deeper investigation regarding the relationship between pre-mRNA and SRm160 in a real time, such as, the state of fluctuations of a molecule of interest (forked DNA or double-labelled RNA) in the presence of SRm160. This chapter is motivated by the previous contribution which identified the state of binding of SRm160 during splicing reaction. The obtained results showed that (1) No FRET signal was appears at the condition supporting the formation of the early complex E. This supports the colocalisation results which indicates a multiple binding of SRm160 molecules alongside the pre-mRNA substrate. This means that the presence of a multiple number of the large molecule (SRm160) prevents the energy transfer. Additionally, this result could be a consequence of limited conformational fluctuations caused by the presence of this rigid molecule. -(2) FRET signal was appeared at the condition supporting the formation of complex A. This finding is also consistent with the colocalisation results at the same condition. Phosphorylation decreased the number of SRm160 proteins. This means that energy transfer between the two flurophores was possible. -(3) The presence of anti-U1 oligo prevented the FRET signal to be detected. This means that SRm160 remained bound to the construct. The persistence of a large molecule affected the FRET signal as identified at the early complex E condition. This finding confirms our conclusion that the binding of SRm160 does not depend on the presence of U1 snRNP.

In all of the proposed methods an automatic system was available to reduce the time required to achieve one experiment. However, to compute results we are looking for, for example, to compute the number of molecules regarding each step, calculation has to done manually for each file of date.

## **ABSTRACT**

---

The manual processes are time consuming and might lead to an unintended errors. To overcome the manual limitations, the proposed methods are then evaluated for data analysis and data documentation.

# Acknowledgments

This thesis would not have been possible without the a great support from my supervisors, colleagues, family and friends. I would like to take this opportunity to express my sincere gratitude to everyone who generously contributed in some way to the work described in this thesis.gave me their support.

First and foremost, I would like to express my deepest sense of gratitude and appreciation to my academic advisors, Professor Ian C Eperon and Dr. Andrew Hudson for accepting me into their group, for their supervision, guidance, patience, engaging me in new ideas, demanding a high quality of work, for being a great friends, for making me appreciate the great experience her in the United Kingdom and for being a great friends. Additionally, I would like to thank my committee members Alison Tyson-Capper and Dr. Andrew Hudson.

I am also grateful for Dr Lucy Eperon for her help and support by sharing her experience to perform better experiments. I thank Dr Olga Makarova, Dr Andrey Revyakin, Dr Dmitry Cherny, Dr Carika Weldon, Dr. Miran Rada, Dr Paul Smith, Dr. Robert Weinmeister, Dr. Li Chen and Dr. Oksana Gonchar for giving me support on my experiments and for their supportive scientific discussions. I would like to thank all members of Lab 3/19, past and present, for their advice and friendship. A deepest thank for Andrew Jobbins, Christian M Lucas for their support they gave me during all these years, enriching my life and offering me help in many aspects inside and outside the

## ACKNOWLEDGMENTS

---

lab for making my graduate life less bearable and full of fun. Gratitude to Dr. Adil Khadidos for his kind support and help.

I would like to extend my thanks to Aditya Chandru and Ghalia Shelmani and many other people in the Henry Wellcome Building for their kind help.

I owe a special thank for Dr. Alaa Khadidos who has supported and constantly encouraged me throughout my PhD and all the inspirations he brings to my life; without whom it would have not been possible.

I also owe much thankfulness to my parents Dr. Abdulaziz Alfi and Afaf Arab for their endless love, encouragements, understanding, and unconditional support during the last four years and at any stage of my life so far. A deep thank and love to my brothers Alaa and Ammar and their wives and to my lovely sister Anan too for their support and encouragement throughout my life.

I would like to dedicate this work to those who left a touch in my life: my relatives and my friends. Their love, constant support and encouragement made me what I am right now.

Last but not least, my successful owe the soul of my grandmother many thanks and appreciation for being praying for me to be what I am now.

# Declarations

I hereby declare that this dissertation entitled *Investigation of the Mechanisms of Action of Exonic splicing Enhancers Using Single Molecule Methods* is an original work and has not been submitted for a degree or diploma or other qualification at any other University

Leicester, United Kingdom.

# Contents

Abstract	i
Acknowledgments	v
Declarations	vii
List of Tables	xv
List of Figures	xvi
Abbreviations	xxiv
<b>Chapter 1 Introduction: <i>Catching a Glimpse of the Life of mRNA</i></b>	<b>1</b>
1.1 RNA Processing and Its Biogenesis Pathway . . . . .	5
1.1.1 Transcription . . . . .	5
1.1.2 The pre-mRNA Processing; 5' Capping, Splicing and 3' Polyadenylation . . . . .	7
1.2 Splicing and the Spliceosomal Machinery . . . . .	10
1.2.1 Messenger RNA . . . . .	10
1.2.2 Pre-mRNA Splicing Reaction . . . . .	11
1.2.3 The Chemical Aspect of Pre-mRNA Splicing . . . . .	13
1.2.4 Spliceosome and Its Association . . . . .	15
1.3 Splice Site Selection and Recognition . . . . .	23

## CONTENTS

---

1.3.1	The 5' Splice Site . . . . .	23
1.3.2	The 3' Splice Site . . . . .	28
1.4	Splicing Regulation . . . . .	30
1.5	The Additional Elements . . . . .	32
1.5.1	Regulatory Elements; Enhancers and Silencer . . . . .	33
1.5.1.1	Splicing Enhancers . . . . .	33
1.5.1.2	Splicing silencers . . . . .	35
1.6	Trans-acting Factors; Additional Proteins Involved in Splicing	38
1.6.1	The Serine-Arginine Proteins; SR Family and SR-related Proteins . . . . .	38
1.6.2	SerineArginine-rich Splicing Co-activator SRm160 and Its Role . . . . .	42
1.7	Single-Molecule Studies . . . . .	46
1.8	Molecular Fluorescence; Theoretical Background . . . . .	48
1.8.1	Absorption and Emission Processes . . . . .	48
1.8.1.1	Absorption Process . . . . .	53
1.8.2	Emission Process . . . . .	55
1.8.3	Fluorophores . . . . .	56
1.9	Total Internal Reflection Fluorescence Microscopy . . . . .	58
1.10	Colocalisation . . . . .	63
1.10.1	Photobleaching . . . . .	64
1.11	Fluorescence Resonance Energy Transfer . . . . .	66
1.12	Single Molecule Experiments on RNA Splicing . . . . .	68
1.13	Contributions of Thesis . . . . .	71
1.14	Thesis Outline . . . . .	72

## CONTENTS

---

<b>Chapter 2</b>	<b>Material and Methods</b>	<b>74</b>
2.1	<i>In vitro</i> Techniques . . . . .	77
2.1.1	Polymerase Chain Reaction (PCR) . . . . .	77
2.1.1.1	The PCR Procedure . . . . .	77
2.1.1.2	DNA Purification . . . . .	77
2.1.2	<i>In vitro</i> Transcription . . . . .	78
2.1.2.1	The procedure of <i>in vitro</i> Hot transcription . . . . .	78
2.1.2.2	The procedure of <i>in vitro</i> Cold Transcription . . . . .	79
2.1.2.3	The procedure of <i>in vitro</i> with 5GMPS CAP and Labelling with Cy5-maleimide . . . . .	79
2.1.3	<i>In vitro</i> Splicing . . . . .	80
2.1.3.1	The procedure of <i>in vitro</i> splicing . . . . .	80
2.1.3.2	Proteinase PK Treatment . . . . .	81
2.1.3.3	Denaturing Polyacrylamide Gel Electrophore- sis . . . . .	81
2.1.4	Analysis of spliceosomal complexes . . . . .	82
2.1.4.1	Native Agarose Gel Electrophoresis . . . . .	82
2.1.5	Western Blot Analysis . . . . .	82
2.1.5.1	The Procedure of WB . . . . .	82
2.1.5.2	SDS-PAGE Gel Electrophoresis . . . . .	83
2.2	Oligonucleotide design . . . . .	84
2.3	Cell Culture and Nuclear Extract Preparation . . . . .	84
2.3.1	Transfection of HEK 293T Cell . . . . .	84
2.3.2	Nuclear Extract Preparation . . . . .	85
2.4	Single Molecule Techniques . . . . .	86
2.4.1	Labelling of the pre-mRNA for colocalisation purposes . . . . .	86

## CONTENTS

---

2.4.2	Chamber Preparation . . . . .	86
2.4.3	Preparation of Samples for the splicing complexes . . .	87
2.4.4	Preparation of Samples for the single molecules experi- ments; colocalisation . . . . .	87
2.4.5	Preparation of Samples for the single molecules experi- ments; FRET . . . . .	88
2.4.6	Samples used for FRET experiments . . . . .	88
<b>Chapter 3 Using of The Microscope: <i>Set up, alignment, data presentation and analysis</i></b>		<b>89</b>
3.1	The Microscope Set up . . . . .	91
3.2	The Two Lasers Fluorescence Microscopy . . . . .	95
3.3	The Camera Vs. the Detection Efficiency . . . . .	96
3.4	The Sample Stability . . . . .	97
3.5	Data Acquisition via LabView . . . . .	98
3.6	Composite Images . . . . .	100
3.7	Spot Detection . . . . .	104
3.8	Inspection of Colocalization . . . . .	104
3.9	The Assignment of Steps . . . . .	108
3.10	Analyzing Data . . . . .	108
3.11	Presentation of Data . . . . .	110
<b>Chapter 4 Optimisation of Condition for Single Molecule exper- iments with SRm160</b>		<b>114</b>
4.1	Labelling of SRm160 in Nuclear Extract . . . . .	116
4.2	Labelling of the Pre-mRNA . . . . .	121
4.3	Tethering the Surface and Modifications . . . . .	121
4.4	Preparation of Splicing Complexes . . . . .	123

## CONTENTS

---

4.4.1	Stalling Splicing at the Early Complex E . . . . .	123
4.4.2	Stalling Splicing at Complex A . . . . .	124
4.4.3	Stalling Splicing at Complex C . . . . .	124
4.5	Pre-mRNA Concentration Required for Single Molecule Experiments . . . . .	125
<b>Chapter 5 Behaviour of SRm160 protein during splicing.</b>		<b>127</b>
5.1	Introduction . . . . .	129
5.1.1	The State of Dimerization . . . . .	130
5.1.2	The Construct . . . . .	131
5.1.3	Control Experiments . . . . .	132
5.2	SRm160 in Different Complexes; <i>the scope from different constructs</i> . . . . .	138
5.2.1	GloC Construct . . . . .	138
5.2.1.1	GloC (Non-Modified Construct) at the Early Complex E (the depletion of ATP) . . . . .	138
5.2.1.2	Globin C Construct at Complex A (the addition of ATP) . . . . .	140
5.2.2	The Interplay of SRm160 Protein With the U1 snRNP and the 5' Splice Site; Globin M . . . . .	143
5.2.3	Intron-less Globin C; mRNA . . . . .	144
5.3	SRm160 and the Splicing Enhancer; SMN2 . . . . .	146
5.4	SRm160 and the Tra2 Site . . . . .	154
5.5	SRm160 Vs. the 5' exon . . . . .	161
<b>Chapter 6 Analysis of Complexes Flexibility Using FRET</b>		<b>173</b>
6.1	Fluorescence Dyes . . . . .	175

## CONTENTS

---

6.2	Fluorescence Resonance Energy Transfer; FRET . . . . .	175
6.2.1	Branched Substrate . . . . .	176
6.2.2	Double-labeled RNA . . . . .	187
<b>Chapter 7</b>	<b>Discussion</b>	<b>196</b>
7.1	The Microscope . . . . .	202
7.2	The Analysis . . . . .	206
7.3	The Preparation . . . . .	208
7.4	The Protein; SRm160 . . . . .	209
7.4.1	The SM-Colocalisation Vs. SRm160 . . . . .	209
7.4.2	The FRET signal Vs. SRm160 . . . . .	216
7.5	Summary . . . . .	217
7.6	Future Work . . . . .	221
<b>Chapter 8</b>	<b>Appendices</b>	<b>223</b>
8.1	Sequences of RNA Constructs . . . . .	224
8.1.1	Glo C . . . . .	224
8.1.2	Globin M . . . . .	224
8.1.3	Globin C GG . . . . .	225
8.1.4	Rabbit $\beta$ -globin Sequence . . . . .	225
8.1.5	CEC . . . . .	226
8.1.6	Primer 16 . . . . .	226
8.1.7	Primer 17 . . . . .	226
8.1.8	SMN2 . . . . .	227
8.1.9	Reverse Tra2 $\beta$ deletion (SMN2) . . . . .	227
8.1.10	Reverse +ESE . . . . .	228
8.1.11	Reverse Tra2 $\beta$ +ESE (SMN2) . . . . .	228

## CONTENTS

---

8.1.12 Frw.primers no (1) . . . . .	228
8.1.13 Frw.primers no (2) . . . . .	228
8.1.14 Frw.primers no (3) . . . . .	228
8.1.15 Frw.primers no (4) . . . . .	229
8.1.16 Frw.primers no (5) . . . . .	229
8.1.17 Frw.primers no (6) . . . . .	229
8.1.18 Frw.primers no (7) . . . . .	229
8.1.19 Frw.primers no (8) . . . . .	229
8.1.20 Frw.primers no (9) . . . . .	230
8.1.21 Frw.primers no (10) . . . . .	230
8.1.22 Promoter 7 . . . . .	230
8.1.23 $\beta$ -globin 5' Cy5 . . . . .	230
8.1.24 $\alpha$ -U1 . . . . .	230
8.1.25 $\alpha$ -U6 . . . . .	231
8.1.26 SRM160 . . . . .	231
8.1.27 Double-Labelled RNA . . . . .	232
<b>References</b>	<b>245</b>

# List of Tables

8.1	Single Molecule Data shows the colocalisation percentages. #M is the the number of identified marker spots, #C is the number of colocalized spots and #CE is the colocalization percentage.	233
8.2	Single Molecule Data shows the number molecules in each bleaching step. #M is the the number of identified marker spots, #C is the number of colocalized spots and #CE is the colocalization percentage. . . . .	237
8.3	Single Molecule Data shows the error bars values. #M is the the number of identified marker spots, #C is the number of colocalized spots and #CE is the colocalization percentage. . .	241

# List of Figures

1.1	Simplified schematic diagram of the key steps behind gene expression . . . . .	6
1.2	The central dogma of molecular biology: the flow of the genetic information mechanism . . . . .	11
1.3	Schematic diagram of the four fundamental sequences required for splicing. . . . .	14
1.4	Simplified schematic diagram of the splicing mechanism . . . .	16
1.5	Major protein composition. . . . .	18
1.6	Assembly and disassembly of the spliceosome. . . . .	20
1.7	Schematic representations of the 5' splice site recognition. . . .	24
1.8	Schematic representations of the 3' splice site recognition. . . .	30
1.9	Schematic representation of the possible distribution of canonical and additional splicing cis-elements. . . . .	32
1.10	The role of SRm160 splicing co-activator in ESE function. . . .	44
1.11	Schematic diagram shows SR domains in SRm160. . . . .	46
1.12	Simplified schematic diagram of Jablonski diagram. . . . .	50
1.13	Simplified schematic diagram of transition energy levels. . . .	51
1.14	Schematic diagram of Stoke's shift. . . . .	52
1.15	Schematic diagram of Beer-Lambert law. . . . .	54

## LIST OF FIGURES

---

1.16 Schematic diagram of refraction at the interface of two different media. . . . .	60
1.17 Schematic diagram of the objective-based illumination microscopy.	62
1.18 Simplified diagram of colocalisation. . . . .	63
1.19 Schematic diagram of the mechanism of FRET. . . . .	67
3.1 A complete view of our home made single molecule microscope.	91
3.2 Our home made single molecule microscope: the main set-up used for single molecule colocalisation measurements. . . . .	93
3.3 Our home made single molecule microscope: the main set-up used for single molecule FRET measurements. . . . .	94
3.4 The Interface window of the LabView. . . . .	101
3.5 Example of a stacked Single Molecule Image. . . . .	102
3.6 Example of a single molecule image images created by calculating the mean or the maximum. . . . .	103
3.7 Example of a single molecule image before and after the enhancement step:illumination by a wavelength of 488 nm. . . . .	105
3.8 Example of a single molecule image before and after the enhancement step:illumination by a wavelength of 632 nm. . . . .	106
3.9 Resulted images from the LabView interface: Prism-based FRET measurements . . . . .	107
3.10 A diagram representation of intensity time traces for GFP with a time resolution of 0.1 s. . . . .	109
3.11 Schematic diagram of the interface of the MATLAB Program.	111
3.12 Exemplary representation of single molecule colocalisation experiment result. . . . .	112

## LIST OF FIGURES

---

4.1	Fluorescence microscopy view of expressed protein. . . . .	117
4.2	Control <i>in vitro</i> splicing reaction: the functionality of the home-made nuclear extract (SRm160) with GloC construct. . . . .	118
4.3	Control <i>in vitro</i> splicing reaction: the functionality of the home-made nuclear extract (SRm160) with different construct. . . . .	119
4.4	Control experiment of western blot. . . . .	120
4.5	Labelling of pre-mRNA via 5'-Cy5-oligonucleotide. . . . .	122
5.1	The state of the SRm160 protein dimerization in the presence of RNaseA. . . . .	131
5.2	Schematic diagram of different constructs of RNA used in single molecule experiment. . . . .	133
5.3	Testing the purity of the slide chamber for single molecule colocalisation experiments. . . . .	134
5.4	Testing the purity and the state of GFP only. . . . .	135
5.5	Testing the proportion of un-colocalised spots (Cy5-RNA). . . . .	136
5.6	Bar chart of the amount of SRm160 present in the nuclear extract: different ratios of SRm160 to commercial NE. . . . .	137
5.7	Schematic diagram of GloC pre-mRNA construct. . . . .	139
5.8	Bar chart of SRm160 binding to GloC pre-mRNA construct: (a) At the early complex E. (b) At complex A. . . . .	141
5.9	Bar chart of SRm160 binding to GloC pre-mRNA construct: (a) At complex A. (b) At complex A in the presence of anti-U1 oligo. (c) At complex A in the presence of anti-U2 oligo. . . . .	142
5.10	Schematic diagram of Glo M construct . . . . .	143

## LIST OF FIGURES

---

5.11	Bar chart of SRm160 binding with Globin M pre-mRNA under different conditions: <b>(a)</b> At the early complex E. <b>(b)</b> At complex A. . . . .	144
5.12	Schematic diagram of spliced GloC RNA . . . . .	145
5.13	Bar chart of SRm160 binding with spliced mRNA (intronless) under different conditions: <b>(a)</b> At the early complex E. <b>(b)</b> At complex A. . . . .	146
5.14	Schematic diagram of SMN2 substrate. . . . .	147
5.15	Bar chart of SRm160 association with SMN2 pre-mRNA under different conditions: <b>(a)</b> and <b>(b)</b> Under condition supporting the formation of the early complex E. <b>(c)</b> and <b>(d)</b> At complex A.	148
5.16	Bar chart of SRm160 association with SMN2 pre-mRNA under different conditions: <b>(a)</b> and <b>(b)</b> At the early complex E. <b>(c)</b> and <b>(d)</b> Using PhosStop in the absence of ATP. . . . .	150
5.17	Bar chart of SRm160 association with SMN2 pre-mRNA under different conditions: <b>(a)</b> and <b>(b)</b> At the early complex E in the presence of anti-U1 oligo. <b>(c)</b> and <b>(d)</b> At the early complex E in the presence of both anti-U1 oligo and phoSTOP. . . . .	151
5.18	Bar chart of SRm160 association with SMN2 pre-mRNA under different conditions: <b>(a)</b> and <b>(b)</b> At complex A in the presence of anti-U1 oligo. <b>(c)</b> and <b>(d)</b> At complex A in the presence of anti-U2 oligo. . . . .	152
5.19	Bar chart of SRm160 association with SMN2 pre-mRNA under different conditions: <b>(a)</b> and <b>(b)</b> At complex A in the presence of both anti-U1 and anti-U2 oligos. . . . .	153

## LIST OF FIGURES

---

5.20	Schematic diagram of SMN2 substrate with and without hTra2 $\beta$ site. . . . .	154
5.21	Bar chart of SRm160 association with SMN2 with and without hTra2 $\beta$ site under different conditions: <b>(a)</b> and <b>(b)</b> At the early complex E. <b>(c)</b> and <b>(d)</b> Using PhosStop in the absence of ATP. . . . .	157
5.22	Bar chart of SRm160 association with SMN2 with and without hTra2 $\beta$ site under different conditions: <b>(a)</b> and <b>(b)</b> At the early complex E. <b>(c)</b> and <b>(d)</b> At complex A. . . . .	158
5.23	Bar chart of SRm160 association with SMN2 with and without hTra2 $\beta$ site under different conditions: <b>(a)</b> and <b>(b)</b> At the early complex E in the presence of anti-U1 oligo. <b>(c)</b> and <b>(d)</b> At the early complex E in the presence of both phoSTOP and anti-U1 oligo. . . . .	159
5.24	Bar chart of SRm160 association with SMN2 with and without hTra2 $\beta$ site under different conditions: <b>(a)</b> and <b>(b)</b> At complex A in the presence of anti-U1 oligo. <b>(c)</b> and <b>(d)</b> At complex A in the presence of anti-U2 oligo. . . . .	160
5.25	The <i>in vitro</i> splicing reaction for a series of GloC 16 17 with different lengths: Samples from (2) to (10) represent the shorter after sample number (1) and so on. . . . .	162
5.26	The <i>in vitro</i> splicing reaction for a series of GloC 16 E3 with different lengths: Samples from (2) to (10) represent the shorter after sample number (1) and so on. . . . .	163
5.27	The <i>in vitro</i> splicing reaction for a series of GloC 16 17 and GloC 16 E3 with different lengths: Samples from (1) to (4) represent the shorter after sample number (1) and so on. . . . .	164

## LIST OF FIGURES

---

5.28	The <i>in vitro</i> splicing reaction for a series of GloC 16 17 and GloC 16 E3 with different lengths: Samples from (5) to (7) represent the shorter after sample number (1) and so on. . . .	165
5.29	The <i>in vitro</i> splicing reaction for a series of GloC 16 17 and GloC 16 E3 with different lengths: Samples from (8) to (10) represent the shorter after sample number (1) and so on. . . .	166
5.30	The <i>in vitro</i> splicing efficiency depends on the mRNA product for a series of different length of GloC substrates. . . . .	167
5.31	The <i>in vitro</i> splicing efficiency depends on the pre-mRNA product at the zero time point for a series of different length of GloC substrates. . . . .	168
5.32	The <i>in vitro</i> splicing efficiency depends on lariat product for a series of different length of GloC substrates. . . . .	169
5.33	The <i>in vitro</i> splicing activity for a series of different length of GloC substrates without the 3' U1 binding site: <b>a</b> Depending on the mRNA results. <b>b</b> Depending on the pre-mRNA results at the zero time point. <b>c</b> Depending on the lariat intron results. .	170
5.34	The <i>in vitro</i> splicing activity for a series of different length of GloC substrates with the 3' U1 binding site: <b>a</b> Depending on the mRNA results. <b>b</b> Depending on the pre-mRNA results at the zero time point. <b>c</b> Depending on the lariat intron results. . . .	171
5.35	The <i>in vitro</i> splicing activity for a series of different length of GloC substrates: A comparison between substrates with and without the 3' U1 binding site. . . . .	172
6.1	Schematic diagram of the DNA branched structure (forked DNA).	176

## LIST OF FIGURES

---

6.2	An exemplary of sample chamber. . . . .	177
6.3	Sample immobilization. . . . .	178
6.4	Testing cover slips (slide chamber) before and after the addition of any solutions (streptavidin and BSA). . . . .	179
6.5	Schematic diagram of Cy3-Sample only: No correction filters were used. . . . .	180
6.6	Schematic diagram of Cy3-Sample only: Correction filters were used. . . . .	181
6.7	Total internal reflection fluorescence microscopy imaging: FRET images from Prism-based TIRF. . . . .	183
6.8	Total internal reflection fluorescence microscopy imaging: FRET time trajectories. . . . .	184
6.9	Schematic diagram of FRET signal of Forked DNA sample: Re- sult 1. . . . .	185
6.10	Schematic diagram of FRET signal of Forked DNA sample: Re- sult 2. . . . .	186
6.11	The structure of the double-labelled RNA construct. . . . .	187
6.12	Schematic diagram of FRET signal of the double-labelled RNA: At the early complex E (result 1). . . . .	189
6.13	Schematic diagram of FRET signal of the double-labelled RNA: At the early complex E (result 2). . . . .	190
6.14	Schematic diagram of FRET signal of the double-labelled RNA: At complex A in the presence of anti-U6 oligo (result 1). . . . .	191
6.15	Schematic diagram of FRET signal of the double-labelled RNA: At complex A in the presence of anti-U6 oligo (result 2). . . . .	192

## LIST OF FIGURES

---

6.16	Schematic diagram of FRET signal of the double-labelled RNA : At complex A in the presence of anti-U1 oligo. . . . .	193
6.17	Comparison diagrams for single molecule signals through different programs. . . . .	195
7.1	The state of SRm160 protein at single molecule level. . . . .	199
7.2	The state of Forked DNA at single molecule level through the fluorescence resonance energy transfer microscopy scope. . . . .	200
7.3	The state of Double-labelled RNA at single molecule level through the fluorescence resonance energy transfer microscopy scope in the presence of our home-made nuclear extract (SRm160). . . . .	201

# Abbreviations

AG	Adenosine-guanine dinucleotide
A	Adenosine
AMPS	Ammonium persulfate
ATP	Adenosine-5-triphosphate
BPS	Branch point Sequence
bp	Base pair
BSA	Bovine Serum Albumin
C	Cytidine
CrPi	Creatine phosphate
Cy5	Indocarbocyanin
C	Degree Celsius
DNA	Deoxyribonucleic Acid
dNTPs	Deoxynucleotide triphosphates (DNA)
dATP	Deoxy adenosine 5 triphosphate
dCTP	Deoxy cytidine 5 triphosphate
dGTP	Deoxy guanosine 5 triphosphate
dTTP	Thymidine 5 triphosphate
DMEM	Dulbeccos Modified Eagles Medium
DMSO	Dimethyl Sulfoxide
DTT	Dithiothreitol

## ABBREVIATIONS

---

ECL	Enhanced chemiluminescence
EColi	Escherichia coli
EDTA	Ethylene diamine tetraacetic acid
EGFP	Enhanced green fluorescent protein
EMCCD	Electron-multiplying charge-coupled device
ESE	Exonic splicing enhancer
ESS	Exonic splicing silencer
EtOH	Ethanol
FBS	Fetal bovine serum
HEK 293T	Human embryonic kidney 293T
Hepes	N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid
hnRNP	Heterogeneous nuclear ribonucleoprotein
ISE	Intronic splicing enhancer
ISS	Intronic splicing silencer
kDa	Kilodalton
K/Glu	Potassium glutamate
LB	Luria broth
LMP	Low melting point

## ABBREVIATIONS

---

mRNA	Messenger RNA
N.E.	Nuclear extract
NP-40	Nonidet
P-40 nt	nucleotide
nts	nucleotides
oligo	oligonucleotide
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PK	Proteinase K
pre-mRNA	precursor messenger RNA
PPT	Polypyrimidine tract
PTB	Polypyrimidine tract binding protein
RNA	Ribonucleic acid
RNase	Ribonuclease
rNTPs	Ribonucleotide triphosphates
RRM	RNA recognition motif
RT	Room Temperature
RON	Receptor dorigine nantais
SDS	Sodium dodecyl sulfate

## ABBREVIATIONS

---

SMA	Spinal muscular atrophy
SMN	Survival of motor neuron
snRNA	Small nuclear ribonucleic acid
snRNP	Small nuclear ribonucleoprotein
SR	Serine-Arginine
SRSF	Serine/Arginine-rich splicing factor
ss	Splice site
TIRF	Total internal reflection
TBE	Tris-borate-EDTA buffer
Tris	Tris(hydroxymethyl)aminomethane
U snRNP	Uridine-rich small nuclear ribonucleoprotein particles
UTP	Uridine 5 triphosphate
U2AF	U2 snRNP auxiliary factor
UV	Ultraviolet
V	Volts
v/v	Volume per volume
W	Watts
w/v	Weight per volume

## ABBREVIATIONS

---

# Chapter 1

## Introduction: *Catching a Glimpse of the Life of mRNA*

Molecular interactions and macromolecular complexes control many cellular processes, such as DNA transcription, RNA splicing and translation. The reaction pathway of these processes passes through a sequence of distinct steps. These steps were identified by the formation of transient complexes unique for each step. For example, the splicing reaction is mediated by the action of an enormous ribonucleoprotein complex termed the spliceosome. Many proteins and several other small nuclear ribonucleoprotein particles are required for this action to be processed.

The accepted view of the cellular processes in eukaryotics is that the genetic information is encoded in the sequence termed deoxyribonucleic acid (DNA), which is stored in the nucleus of the cell. The DNA is transcribed first into messenger ribonucleic acid (mRNA). The resulting mRNA then gets transported from the nucleus into the cytoplasm, where it is translated into proteins.

Overall, from the transcription of the genetic information stored

---

in the DNA to protein synthesis, many regulatory processes are introduced. This regulatory system controls the identity and abundance of RNAs and proteins. In addition, it contributes to forming the final repertoire of molecules present in the cell. In the present chapter, an overview of the key steps behind this system, with particular emphasis on RNA splicing, is provided.

# Contents

1.1	RNA Processing and Its Biogenesis Pathway . . . . .	5
1.1.1	Transcription . . . . .	5
1.1.2	The pre-mRNA Processing; 5' Capping, Splicing and 3' Polyadenylation . . . . .	7
1.2	Splicing and the Spliceosomal Machinery . . . . .	10
1.2.1	Messenger RNA . . . . .	10
1.2.2	Pre-mRNA Splicing Reaction . . . . .	11
1.2.3	The Chemical Aspect of Pre-mRNA Splicing . . . . .	13
1.2.4	Spliceosome and Its Association . . . . .	15
1.3	Splice Site Selection and Recognition . . . . .	23
1.3.1	The 5' Splice Site . . . . .	23
1.3.2	The 3' Splice Site . . . . .	28
1.4	Splicing Regulation . . . . .	30
1.5	The Additional Elements . . . . .	32
1.5.1	Regulatory Elements; Enhancers and Silencer . . . . .	33
1.5.1.1	Splicing Enhancers . . . . .	33
1.5.1.2	Splicing silencers . . . . .	35
1.6	Trans-acting Factors; Additional Proteins Involved in Splicing	38

## CONTENTS

---

1.6.1	The Serine-Arginine Proteins; SR Family and SR-related Proteins . . . . .	38
1.6.2	SerineArginine-rich Splicing Co-activator SRm160 and Its Role . . . . .	42
1.7	Single-Molecule Studies . . . . .	46
1.8	Molecular Fluorescence; Theoretical Background . . . . .	48
1.8.1	Absorption and Emission Processes . . . . .	48
1.8.1.1	Absorption Process . . . . .	53
1.8.2	Emission Process . . . . .	55
1.8.3	Fluorophores . . . . .	56
1.9	Total Internal Reflection Fluorescence Microscopy . . . . .	58
1.10	Colocalisation . . . . .	63
1.10.1	Photobleaching . . . . .	64
1.11	Fluorescence Resonance Energy Transfer . . . . .	66
1.12	Single Molecule Experiments on RNA Splicing . . . . .	68
1.13	Contributions of Thesis . . . . .	71
1.14	Thesis Outline . . . . .	72

### 1.1 RNA Processing and Its Biogenesis Pathway

The sequential steps of mRNA biogenesis are not independent (see figure 1.1) and are linked to each other. Functional connections between the protein factors that are responsible for the different steps in the gene expression pathway have been reported. Interactions between the various machinery have also been identified. Regulation of the pathway from gene to protein is also controlled at different stages [Moore, 2005; Orphanides and Reinberg, 2002]. Figure 1.1 simplifies the biogenesis pathway of the RNA. mRNA molecules undergo various modifications before they are exported from the nucleus to the cytosol. These modifications imply various routes. These routes include the addition of a 5' cap (5' end processing), splicing (removal of introns) and the polyadenylation of the 3' end (3' end processing) [Darnell, 2013]. The following sections shed light on the fundamental processes of RNA and its modifications.

#### 1.1.1 Transcription

Eukaryotic genes are characterised by the presence of sequences that are included or excluded called exons and introns, respectively. Transcription can be initiated and catalysed by three different enzymes, i.e. RNA polymerases I, II and III. RNA polymerase II (RNA pol II) is commonly identified to be the most notable player in the transcription reaction. On the other hand, RNA polymerase I (RNA pol I) and RNA polymerase III (RNA pol III) are involved in the transcription of ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) [Orphanides and Reinberg, 2002; Paule and White, 2000; Woychik and Hampsey, 2002]. Overall, transcription via RNA polymerase involves mul-

## 1.1 RNA Processing and Its Biogenesis Pathway

---

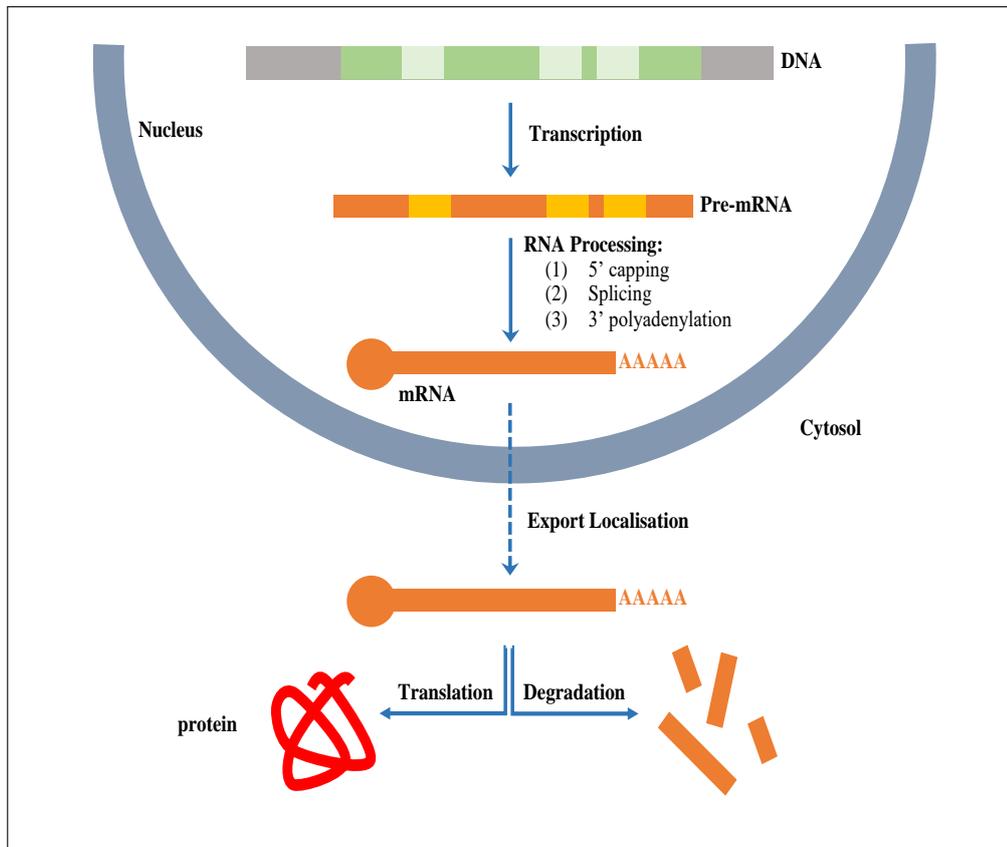


Figure 1.1: Simplified schematic diagram of the key steps behind gene expression. RNA expression initiates with the transcription of DNA. Following several processing steps, for instance, the transcription products are further transformed into mature mRNAs that can then be exported to the cytosol. Once in the cytosol, mRNAs can be recognized by ribosomes and translated into proteins or will be eventually degraded.

## 1.1 RNA Processing and Its Biogenesis Pathway

---

multiple steps. It starts with the binding of several proteins, commonly regarded as transcription factors (TFs) to a promoter region. This region is located upstream of the gene and acts as a regulatory part [Fuda et al., 2009]. These proteins enable the assembly of the polymerase and the formation of the transcription initiation complex. They also participate in the recognition of the majority of promoters. Contrariwise, other specific TFs also exist, which can modulate the fate of the reaction by binding to DNA regions that promote (enhancers) or inhibit (silencers) polymerase assembly [Vaquerizas et al., 2009]. Further conformational rearrangements are also involved. Following the previous steps, RNA polymerase is released from the complex of proteins and leaves the promoter (promoter clearance). By this step, RNA polymerase enters the elongation stage [Kwak and Lis, 2013]. Ultimately, the polymerase transcribes through the cleavage and polyadenylation signals that mark the end of the gene, and it is released from the DNA template [Kuehner et al., 2011]. The resulting pre-mRNA then undergoes further reactions that involve multiple modifications in the nucleus to form a mature messenger RNA (mRNA). The processing includes capping, splicing and 3' end processing. Thereafter, mRNA is transported from the nucleus to the cytoplasm for protein synthesis via translation [Soller, 2006].

### 1.1.2 The pre-mRNA Processing; 5' Capping, Splicing and 3' Polyadenylation

After the RNA polymerase has entered the elongation phase, a methylated guanine nucleotide (cap) and N7-methyl GMP (guanosine monophosphate) is added to the 5' end of the pre-mRNA via 5' - 5' triphosphate linkage [McCracken et al., 1997a,b] through a reaction termed capping. Such process

## 1.1 RNA Processing and Its Biogenesis Pathway

---

delimits the 5' end of transcripts and enables the distinction of mRNAs from other RNA species (e.g. RNA pol I and III produce uncapped RNAs). It also protects the RNA molecule from degradation by protecting the mRNA from 5' - 3' exonucleases. Capping also aids in the initiation of translation by promoting the binding of the ribosomal subunits to the mRNA [Proudfoot et al., 2002]. It was also demonstrated that capping enhances splicing 3' end and export processing [Flaherty et al., 1997; Fortes et al., 2000; Izaurralde et al., 1994; Lewis et al., 1996]. After that, pre-mRNA undergoes a process termed as splicing. Splicing, however, is a much more complex reaction. Splicing is a process whereby some regions of the pre-mRNA (introns) are removed, and other that contains the necessary information for protein synthesis (exons) are brought or spliced together to produce mRNA [Black, 2003]. The mechanism of splicing and its regulation is detailed later in this chapter.

Similarly, the 3'-end of mRNAs is also modified. Such modification involves the addition of a polyadenosine (polyA) tail. This process, named as polyadenylation, serves the function of extending the mRNA half-life [Proudfoot et al., 2002; Zhao et al., 1999].

Once a mature RNA (mRNA) molecule is synthesized via splicing and before being translated, it is exported from the nucleus to the cytoplasm through the nuclear pore. The export process is linked to strict quality control mechanisms that ensure that immature RNAs remain in the nucleus [Porrúa and Libri, 2013]. Those mechanisms rely on the recognition of protein complexes that accompany the RNA molecules (i.e. RNA-binding proteins; RBPs), which act as markers of the completion status of the steps mentioned in the previous section. For example, the cap-binding and polyA binding-complexes act as indicators of successful capping and polyadenylation reactions, respectively,

## 1.1 RNA Processing and Its Biogenesis Pathway

---

and other protein complexes mark the end of the splicing in a similar fashion (i.e. the exon-junction complex (EJC)). Conversely, the presence of RBPs involved in the execution of each of these steps marks the mRNA molecule as immature, such that, RBPs prevents the export of mRNA.

Unprocessed mRNAs, together with the remainder from the transcription and splicing reactions, will be degraded by a large complex of RNA exonucleases termed the exosome [Pérez-Ortín et al., 2013]. In the cases when they are erroneously exported, or when intact mRNAs become damaged in the cytosol, further control mechanisms prevent their translation. Most of these are intrinsic to the steps required for the initiation of protein synthesis. However, another separate surveillance system also exists. This system actively seeks aberrant mRNAs for degradation, before efficient translation occurs. This system is referred as nonsense-mediated decay (NMD) and specifically targets the presence of premature stop codons in the transcript, which might arise from errors in the splicing reaction [Pérez-Ortín et al., 2013]. During NMD, the first round of translation starts when the 5'-end of the mRNA emerges from the nuclear pore, during which the exon-junction complexes that surround each splice site are detached from the mRNA. Under the presence of nonsense codons, the mRNA remains bound to such complexes and is rapidly degraded.

Following the export process, mRNAs are then localised within the cytosol according to the signals encoded in their 3'-UTR regions, and are eventually recognised by ribosomes and translated [Alberts et al., 2002]. Such binding of ribosomes to the mRNAs is in direct competition with mRNA decay, a process that starts when transcripts are exported into the cytosol. Once the polyA tail reaches a critical length, mRNAs are ultimately degraded. This degradation is either through the continuation of the digestion from the 3'-

## 1.2 Splicing and the Spliceosomal Machinery

---

end or through the removal of the 5'-cap (i.e. decapping) and subsequent 5' to 3' decay [Schoenberg and Maquat, 2012].

## 1.2 Splicing and the Spliceosomal Machinery

Most genes are interrupted by non-coding sequences termed introns. Thus, to generate a functional message from the DNA, introns must be removed from the resulted RNA and coding sequences (exons) are spliced together. The coming sections shed light with more details about splicing and its mechanisms.

### 1.2.1 Messenger RNA

Messenger RNA (mRNA) was discovered in 1961. Various discoveries on understanding the molecularity of DNA and how information within the genetic code is translated were identified. During this time, Crick defined the central dogma of molecular biology, which provided a framework for how genetic information flows between the key biological molecules; DNA, RNA and protein [Crick et al., 1970; Watson and Crick, 1958], see figure 1.2. The flow of the genetic information is sequential. Therefore, this information is transferred from DNA to RNA to protein. Such that Crick's model (the central dogma) was reformulated to include the discovery of reverse transcription [Baltimore, 1970; Crick et al., 1970; Mizutani and Temin, 1970; Temin and Mizutani, 2010]. A major surprise came in 1977 from the laboratories of Phillip Sharp and Richard Roberts. Their studies had concentrated on how DNA is transcribed into mRNA and the relationship between them.

Additional studies demonstrated that eukaryotic genes also have coding sequences that have been split by non-coding sequences [Brack and Tonegawa,

## 1.2 Splicing and the Spliceosomal Machinery

---

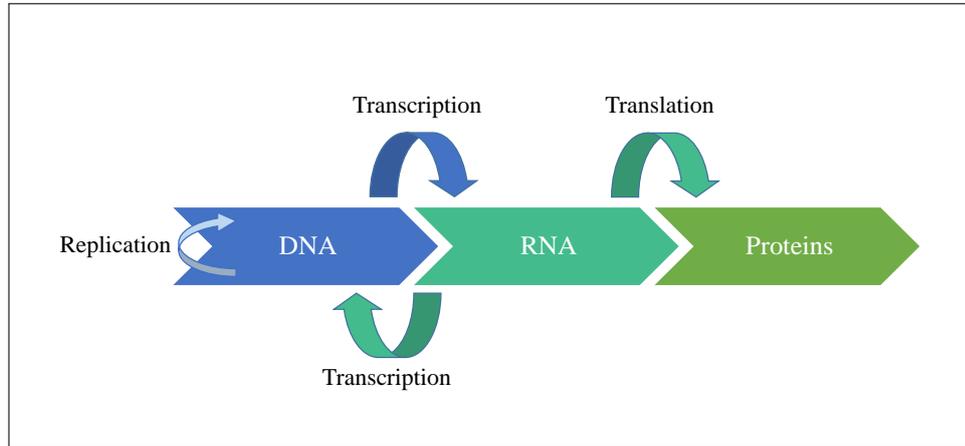


Figure 1.2: Simplified schematic diagram of the central dogma of molecular biology. Each arrow represents the flow of the genetic information mechanism.

1977; Jeffreys and Flavell, 1977]. These intervening, non-coding sequences were defined later as introns, and the corresponding coding sequences being spliced together as exons [Gilbert, 1978]. Since the initial discovery of pre-mRNA splicing, further research has shed light on the key factors of splicing, for example, the spliceosome composition, its machinery, how this complex regulates the splicing of fragments of pre-mRNA, and how aberrant regulation of splicing can lead to various diseases. A brief description of spliceosome and RNA splicing is listed in the following sections.

### 1.2.2 Pre-mRNA Splicing Reaction

The initial transcription of many genes in eukaryotes results in pre-mRNA containing non-coding sequences (introns) that are removed from pre-mRNA, leaving only parts, or coding sequences (exons) within the spliced RNA. The resulted RNA (mature mRNA) contains only exons. This process is known as RNA splicing. The mechanism of splicing and its regulation is explained later in this chapter.

## 1.2 Splicing and the Spliceosomal Machinery

---

The splicing reaction is a major step in gene expression in eukaryotes. This process takes place within two primary chemical steps. In addition, the splicing reaction occurs in the presence of a complex machine called a spliceosome. The first system used for studying mRNA splicing and the structure of the pre-mRNA was the late stage of adenovirus infection in mammalian cells. The presence of introns was first described in the mRNA segment of adenovirus, coding for the Hexon polypeptide, the major virion structural protein [Berget et al., 1977]. In this system, different RNA called (mosaics) were identified, which contained sequences from non-contiguous sites in the viral genome. Studying the process before infection identified that the long RNA contained the sequence of the late RNA in addition to the sequences that were called intervening sequences; introns. After the discovery in adenovirus of the presence of long sequences, the introns were found in other viral and eukaryotic genes such as haemoglobin and immunoglobulin [Breathnach et al., 1978]. After the characterization of the RNA splicing process in viruses, the presence of introns was also reported in eukaryotic genes.

The maturation of RNA, which follows pre-mRNA transcription, is required not only for the splicing process but also for other steps such as capping, polyadenylation and RNA editing. Some studies have outlined functional relationships between these maturation processes and each of them (except the editing process) can occur co-transcriptionally. Many findings point out that during mRNA biogenesis a complex network of functional interactions is formed. Different processing machineries exist and are responsible for capping, splicing, polyadenylation, modification and transport of mRNAs. They can interact with the elongating RNA polymerase II, suggesting that all these processes are physically and functionally intertwined. Consequently, RNA

## 1.2 Splicing and the Spliceosomal Machinery

---

transcription and processing seem to be functionally combined to maximise efficiency and to extend the possibility of regulation [Bentley, 2002; Kornblihtt et al., 2004; Maniatis and Reed, 2002; Proudfoot et al., 2002].

In fact, the process of splicing occurs in all eukaryotic species, and the chemistry of splicing is highly conserved from yeast to human, although yeast has a few short introns and human have multiple long introns. The high prevalence of introns in humans contributes to proteomic diversity via a process called alternative splicing, a mechanism that might allow the use of different or altered coding parts, resulting in a wide variety of different mRNA products that contribute to the formation of distinct proteins and therefore proteins from one apparent gene. A change in alternative splicing can be related to human disease. Unfortunately, most changes caused by alternative splicing are hard to detect because of the complexity of the splicing process [Kelemen et al., 2013]. In fact, this is one of the outstanding differences between species. It was found that there are significant variations in alternative splicing complexity in different species, with the highest complexity found in primates [Barbosa-Morais et al., 2012].

### 1.2.3 The Chemical Aspect of Pre-mRNA Splicing

The process of splicing is complex. This process requires multiple different proteins and other several components (small RNAs) to form the spliceosome complex. The spliceosome involves five small nuclear ribonucleoprotein particles, (U1, U2, U4, U5 and U6), and 170 other associated proteins [Wahl et al., 2009]. The spliceosome complex is confined to the nucleus [Steitz et al., 2008]. The splicing process requires four fundamental sequences in the pre-mRNA (Figure1.3). The junctions of introns and exons are denoted as the 5'

## 1.2 Splicing and the Spliceosomal Machinery

---

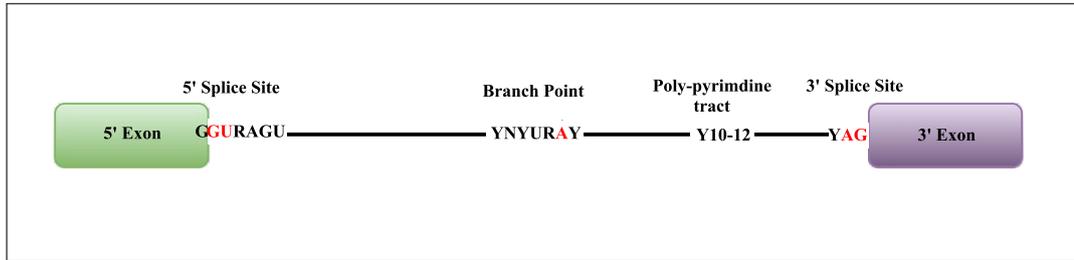


Figure 1.3: Schematic diagram of the four fundamental sequences required for splicing. From left; the 5' exon (the green box), the 5' splice site, the branch point, the poly-pyrimidine tract, the 3' splice site and the 3' exon (the purple box).

splice site and 3' splice site, respectively. These sequences are conserved sequences. Their sequences are G—GURAGU and YAG [Sheth et al., 2006]. The polypyrimidine tract (PPT) is located directly upstream of the 3' splice site (3'ss). Between this tract and the 5' splice site (5'ss) is the branch point (BP) with a sequence around the important adenine [Reed and Maniatis, 1988].

The splicing reaction at its heart takes place through two consecutive chemical steps (trans-esterification reaction): the formation of the lariat intron and exon ligation. Via this reaction and as the first step of the splicing reaction, the 5' exon is displaced through nucleophilic attack of the phosphodiester upstream of the guanine and uracil at the 5' splice site by the 2'-hydroxyl of the adenosine. This results in the free 5' exon and a lariat intermediate attached to the 3' exon sequences (Figure1.4). In this reaction, the ester bond between the 5'-phosphorus of the intron and the 3'-oxygen of exon (1) is exchanged for an ester bond with the 2'-oxygen of the branch-site (A) residue. In the second reaction, the 3'-hydroxyl of the 5' exon attacks the phosphodiester bond following the adenine and guanine at the 3' splice site, separating the exon on this side from the intron and splicing the 5' splice site and 3' splice site of the exons together. By this step, the intron lariat is now fully released, leaving

## 1.2 Splicing and the Spliceosomal Machinery

---

the exons in the mature mRNA. In this reaction, the ester bond between the 5'-phosphorus of exon (2) and the 3'-oxygen of the intron is exchanged for an ester bond with the 3'-oxygen of exon (1), releasing the intron as a lariat structure and joining the two exons [Konarska et al., 1985; Moore and Sharp, 1993; Ruskin et al., 1984; Schellenberg et al., 2008; Umen and Guthrie, 1995]. The outstanding feature of the steps of the splicing reaction is not the underlying chemistry. The recognition of the correct splice sites and bringing them close enough together for the two catalytic steps are the challenges. Despite these, the whole process is highly efficient, which is achieved by the highly dynamic macromolecular machinery consisting of a large number of subunits and factors, the spliceosome. The following section details the spliceosome and its machinery.

### 1.2.4 Spliceosome and Its Association

In eukaryotes, the splicing reaction is carried out through the spliceosomal pathway, whereby a large complex of proteins and RNAs co-ordinates the process of intron removal that harbour consensus splice site sequences [Matera and Wang, 2014] and exons ligation. Such a complex is known as the spliceosome, which is specifically involved in the splicing events and has been designated as one of the most complicated machineries in the cell [Nilsen, 2003].

#### a. The Core of the Spliceosome

The accuracy of pre-mRNA splicing requires the assembly of a large ribonucleoprotein (RNP) complex known as the spliceosome, which is considered as "the most complex macromolecular machine in the cell" [Nilsen, 2003]. It has a molecular mass of  $\sim 2.7$  MDa [Wahl et al., 2009].

## 1.2 Splicing and the Spliceosomal Machinery

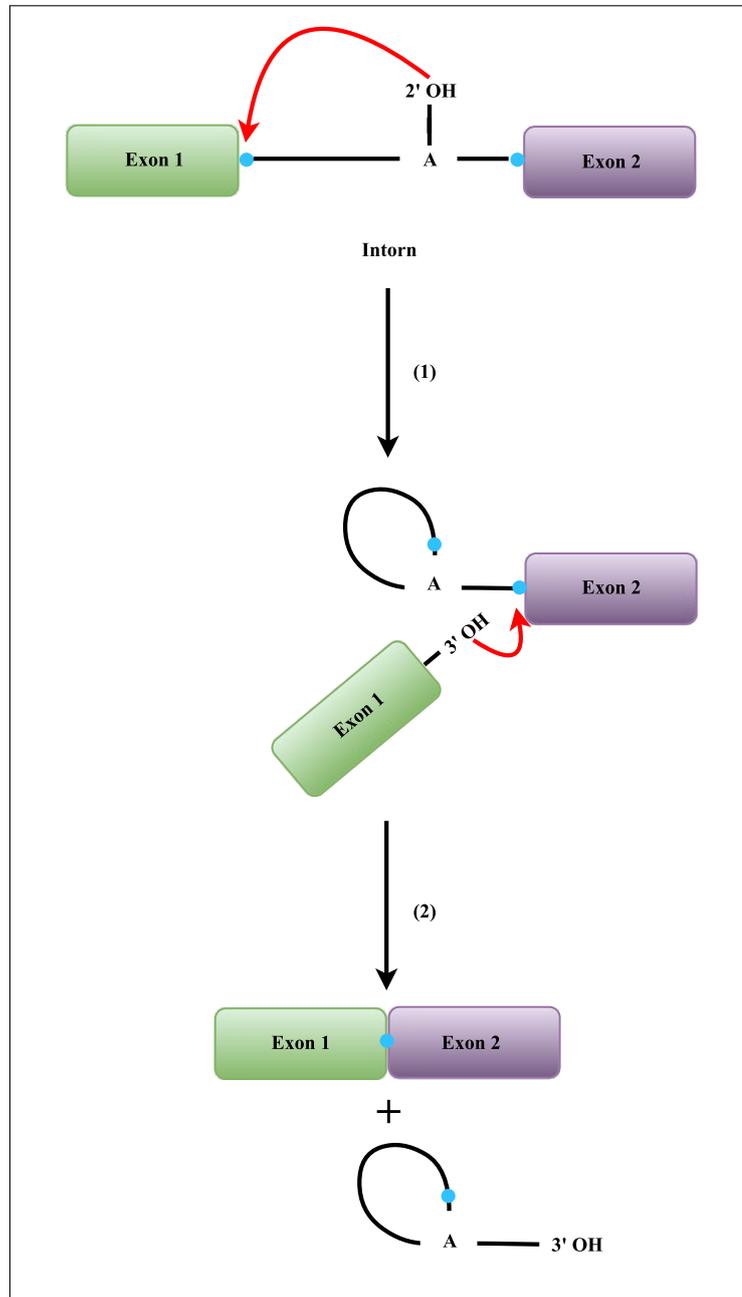


Figure 1.4: Simplified schematic diagram of the splicing mechanism. From top, the solid black line represents a single intron flanked between two exons (exon 1 and exon 2). Green box represents the 5'-end exon and the purple box represents the 3'-end exon. The first (1) and the second (2) steps of splicing involve nucleophilic attacks (red arrows) on phosphodiester bonds (blue dot) by the 2'-OH of the branch point adenosine (A) followed by an attack of the 3'-OH of the free 5' exon at the 3' exon. The ligated exons (spliced RNA) and the lariat product are shown on the bottom.

## 1.2 Splicing and the Spliceosomal Machinery

---

The core components of the spliceosome consist of small nuclear ribonucleoprotein particles (snRNPs). Each snRNP is a complex composed of five snRNAs and other associated protein subunits [Mamatis and Reed, 1987; Nilsen, 2003]. SnRNPs formed by snRNAs (U-rich groups) bind with a set of Sm proteins (U1, U2, U4 and U5) or LSM proteins (U6) [Will and Lührmann, 2011]. The U4, U5 and U6 snRNPs form a tri-snRNP by partial snRNA base pairing as U4 and U6 form a di-snRNP (Figure 1.5). These snRNPs are essential for splicing [Black et al., 1985; Black and Steitz, 1986; Chabot et al., 1985; Krainer and Maniatis, 1985; Krämer et al., 1984; Padgett et al., 1983; Rogers and Wall, 1980].

The spliceosome has to be assembled first before it has its splicing activity. Such assembly, for each interaction with the pre-mRNA, is done as ultimately it is reassembled. The assembly and disassembly of the spliceosome is a dynamic process. This dynamic process undergoes essential compositional and structural rearrangements in sequential steps [Hoskins and Moore, 2012] to catalyse the two transesterification reactions required to excise introns and ligate exons [Wahl et al., 2009]. These rearrangements are assisted by the actions of multiple components, such as protein kinases, phosphatases and RNA helicases [Matlin and Moore, 2006].

### b. Spliceosome Assembly

The model of spliceosome assembly has been characterised using *in vitro* approaches [Matlin and Moore, 2006; Wahl et al., 2009]. A large body of evidence supports a model where the spliceosome assembly is a step-wise process involving recruitment of snRNPs and proteins to the pre-mRNA

## 1.2 Splicing and the Spliceosomal Machinery

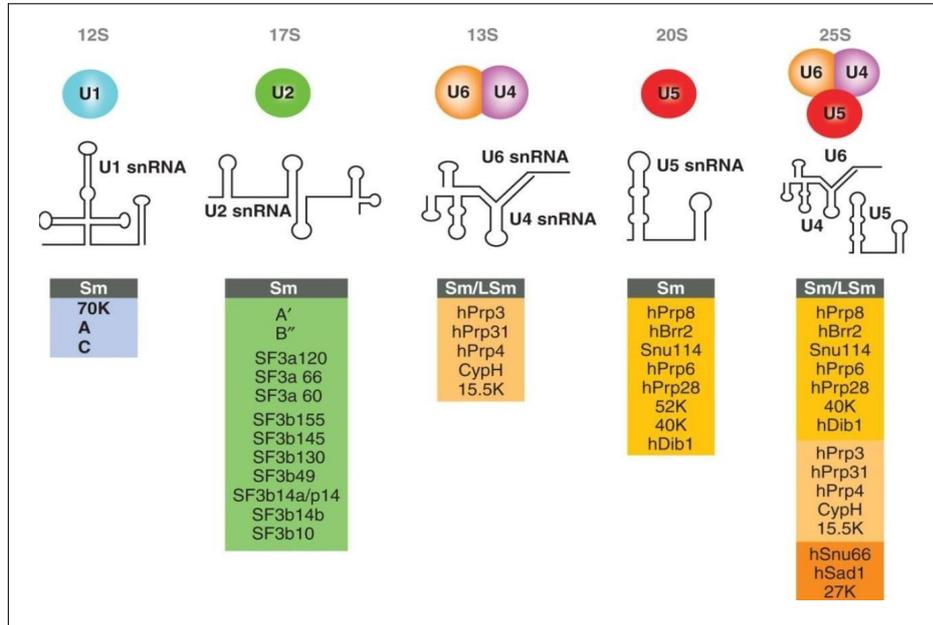


Figure 1.5: Major protein composition. From top; the coloured circles represent all main snRNP involved in the spliceosome formation. The following row represents the structure of each snRNP. This figure is adapted from [Will and Lührmann, 2011].

and dynamic rearrangements of RNA-RNA, RNA-protein and protein-protein interactions (Figure 1.6) [Matlin and Moore, 2006; Zhang and Rosbash, 1999]. Their differential migration has identified the intermediates of this step-wise assembly during native gel electrophoresis, glycerol gradient centrifugation and affinity purification of these complexes during *in vitro* splicing catalysis [Bindereif and Green, 1987; Grabowski and Sharp, 1986; Konarska and Sharp, 1986; Lamond et al., 1988]. More recently, a fluorescence-based method and chromatin immunoprecipitation experiments have provided additional support for the assembly model of the spliceosome [Görnemann et al., 2005; Hoskins et al., 2011a,b; Lacadie and Rosbash, 2005]. The assembly of the spliceosome is divided into distinct stages, termed complexes, starting with the early complex (E),

## 1.2 Splicing and the Spliceosomal Machinery

---

followed by complexes A, B, B\* (or Bact), C and a post-spliceosomal complex (Figure 1.6) [Rino et al., 2007]. It has been supported that the earliest step in spliceosome assembly is the formation of the early complex (E). This formation occurs in an (ATP)-independent manner [Jamison and Garcia-Blanco, 1992; Michaud and Reed, 1991, 1993; Reed, 1990]. It also involves the association of U1 snRNP and other additional protein factors [Bennett et al., 1992; Michaud and Reed, 1991, 1993], see figure 1.6.

In the beginning, U1 snRNP recognises the 5' splice site (5'ss). Recognition of the (5'ss) involves base pairing between the 5'-terminus of U1 snRNA and the pre-mRNA [Juan et al., 2014]. This binding is further stabilised by serine/ arginine-rich proteins (SR) and proteins of the U1 snRNP. Further interactions are also involved. Splicing factor 1 (SF1) binds the branch point. In addition, there is binding of the U2 auxiliary factor; the 65-kDa and 35-kDa subunits (U2AF65 and U2AF35) to the (PPT) and the (AG) sequence at the 3'-splice site, respectively [Berglund et al., 1998; Krämer and Utans, 1991; Wu et al., 1999; Zamore et al., 1992]. These interactions at both the 3'- and the 5'- splice sites arrange the initial recognition of the intron, which serve as platforms for additional interactions with other snRNPs and auxiliary proteins that will occur in subsequent steps. All of these interactions are supported by additional proteins, including those with domains composed of alternating arginine and serine dipeptide repeats (RS domains; described in more detail later in this chapter) [Lin and Fu, 2006]. After the assembly of the early complex (E), complex (A) forms. This complex is

## 1.2 Splicing and the Spliceosomal Machinery

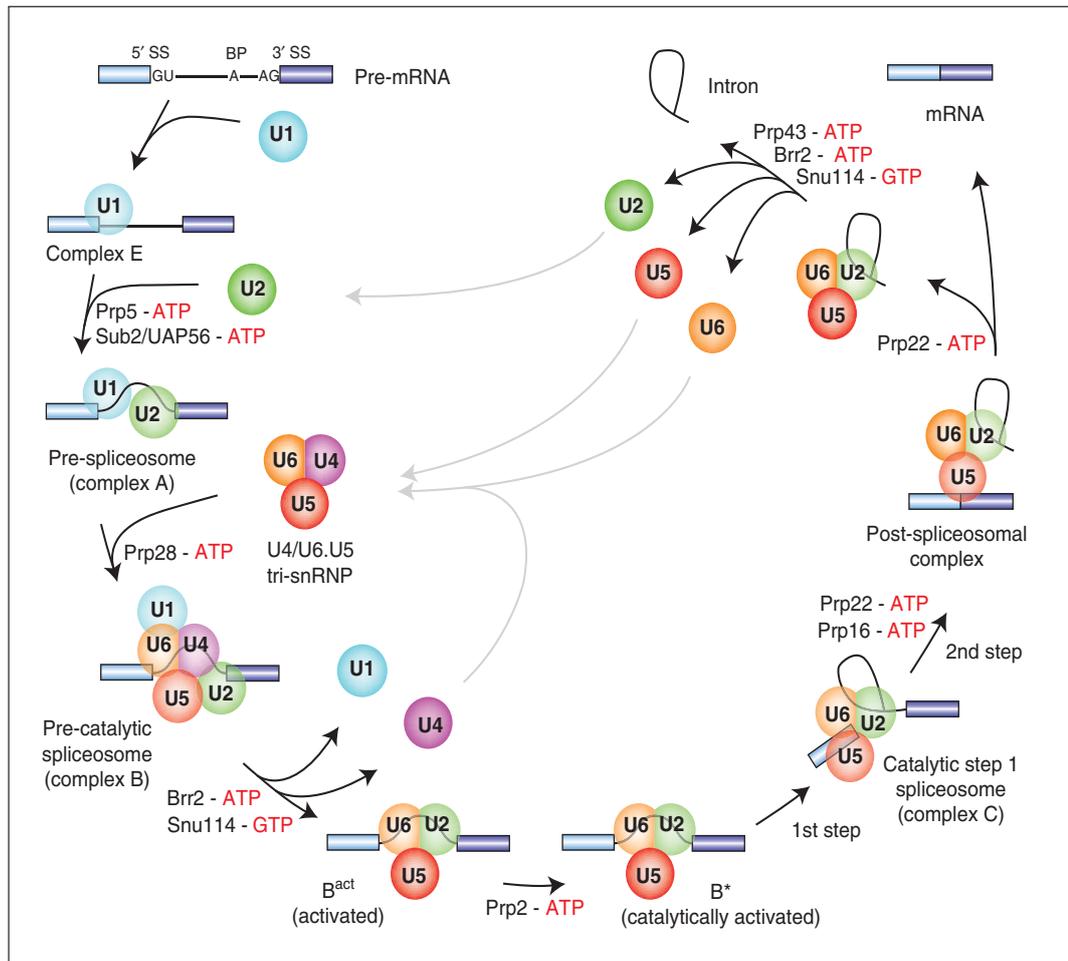


Figure 1.6: Assembly and disassembly of the spliceosome. Spliceosomal complexes form in an ordered way with the two main catalytic reactions occurring before and after complex C. This process removes intron and ligates the exons together. Afterwards the spliceosome is disassembled into its building blocks. This Figure is modified from [Will and Lührmann, 2011]

## 1.2 Splicing and the Spliceosomal Machinery

---

characterised by the binding of the U2 snRNP to the branch site [Wu and Manley, 1989]. The association between U2 snRNA and the pre-mRNA involves the replacement of (SF1) at the branch point via the U2 snRNP [Liu et al., 2001; Rutz and Séraphin, 1999]. This U2 binding requires (ATP), and it is further stabilized by other proteins, most notably U2AF65 [Barabino et al., 1990; Chiara et al., 1996; Gozani et al., 1998; MacMillan et al., 1994; Query et al., 1994; Shen and Green, 2004; Shen et al., 2004; Valcárcel et al., 1996; Valcárcel and Green, 1996]. After complex (A) assembly, a pre-associated tri-snRNP complex composed of U4/ U6 and U5 snRNAs joins the spliceosome and interacts with the pre-mRNA along with other proteins to form complex (B) [Bindereif and Green, 1987; Cheng and Abelson, 1987; Konarska and Sharp, 1987], see figure (1.6). The spliceosome now contains all the necessary components [Deckert et al., 2006] but is not yet active as it requires major conformational rearrangements to form the catalytically active spliceosome. The remodelling interactions during spliceosome activation include multiple steps: (1) the unwinding of the base-paired U4 and U6 snRNAs. (2) Dissociation of U4 snRNP to allow base pairing of U6 snRNA with the (5'ss), (3) replacement of U1 snRNA at the 5' splice site with U6 snRNA, (4) dissociation of the U1 snRNP, (5) base-pair formation between U2 and U6 snRNAs, (6) release of the U1 and U4 snRNPs and (7) rearrangement of interactions between U2 and U6 snRNA. After further changes in the protein composition, complex B becomes active as the complex B\* [Bessonov et al., 2010]. By this stage, the spliceosome is ready to carry out the first transesterification reaction where the sugar backbone phosphodiester bond at the 5' splice site is attacked by the

## 1.2 Splicing and the Spliceosomal Machinery

---

2'-hydroxyl group of the branch point adenosine and a free 3'-hydroxyl group is generated at the end of the upstream exon. This process leads to link the 5'-end of the intron to the branch point, forming the intron lariat, changing over to complex (C). At this stage, the spliceosome is poised to complete the second catalytic step, where the 3'-hydroxyl group of the upstream exon attacks the phosphodiester bond at the 3'-splice site, ligating the two exons. Following these two transesterification reactions of splicing, the products (both intron lariat and exons) are subsequently released, and the components of the spliceosome are recycled [Staley and Guthrie, 1999]. However, catalysis cannot occur without additional conformational rearrangements which facilitated by different factors. In particular, the lariat intermediate must be displaced from the active site allowing the 3'-splice site to be positioned in close proximity to the 3'-hydroxyl group of the upstream exon [Rhode et al., 2006; Umen and Guthrie, 1995]. Once the second transesterification reaction is completed, a multi protein assembly known as the exon-junction complex is deposited on the transcript in the vicinity of the site where splicing has occurred [Tange et al., 2004].

Overall, interactions between the U2, U5 and U6 snRNPs and the pre-mRNA are destabilised at this point, allowing the snRNPs to be recycled and assembled in conformations amenable for another round of splicing. The excised intron lariat is debranched and digested to free nucleotides that are also recycled for future use.

### 1.3 Splice Site Selection and Recognition

---

## 1.3 Splice Site Selection and Recognition

The splicing process is directed by the presence of specific sequences at the exon/ intron boundaries known as 5' and 3' splice sites [Burge et al., 1999]. The recognition of these sites is the key and the challenge of splicing to be accurate. It was identified that the cores of these sites are conserved sequences; the dinucleotide (GU) is the core sequence for the 5' splice site, the dinucleotide (AG) is the core sequence for the 3' splice site and the adenosine (A) is the core sequence for the branch point (BP) [Aebi et al., 1987].

### 1.3.1 The 5' Splice Site

Identifying the exon/ intron junction is a key process in splicing. This process is accomplished by the selection of the 5' splice site (Figure 1.7). As the average intron is much longer than the average exon (International Human Genome Sequencing Consortium, 2001) and a multitude of 5' splice site sequences exist, numerous possible 5' splice sites are present. It has been demonstrated that eukaryotes contain nine partially conserved nucleotides, MAG/ GURAGU (M indicates A or C, R indicates purines and the slash the exon-intron boundary) at the exon-intron junction [Roca and Krainer, 2009]. The (GU) dinucleotide is the core of the 5' splice site consensus sequence and is also universally conserved. Mutations in one of these two nucleotides completely inhibit the splicing process [Juan et al., 2014]. [Zhuang and Weiner, 1986] also demonstrated the importance of those dinucleotides in the recognition of 5' splice site with mutations in the RNA part of the U1 snRNP responsible for sequence recognition, leading to different recognised splice site sequences.

As the exon downstream of the intron is cleaved between the guanine

### 1.3 Splice Site Selection and Recognition

---

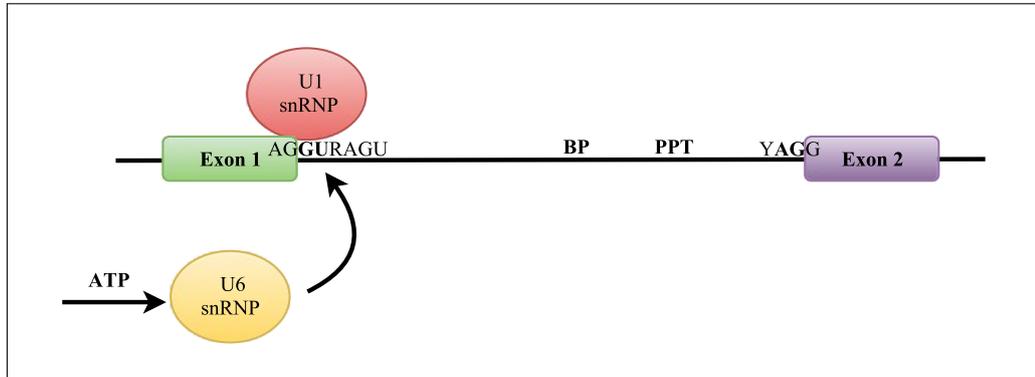


Figure 1.7: Schematic representations of the 5' splice site recognition. The figure represents the recognition of the 5' splice site by the U1 snRNP during the early assembly of the spliceosome followed by the recognition by the U6 snRNP in the presence of ATP. The green and the purple boxes represent the exons, and the solid black line represents the intron. The red and the yellow circles represent the U1 snRNP and the U6 snRNP, respectively.

(G) and uracil (U) that are base-paired to the U1 snRNP, the binding of the U1 snRNP itself plays a major role in selecting the correct splice site. It was demonstrated that the U1 snRNP is the first snRNP to base pair to candidate 5' splice sites both *in vivo* and *in vitro* [Görnemann et al., 2005; Huranová et al., 2010; Reed, 1990; Roca et al., 2013; Wassarman and Steitz, 1992].

It has also been shown that the recognition of the 5' splice site by the U1 snRNA is accomplished during the early assembly of the spliceosome machinery [Horowitz and Krainer, 1994; Siliciano and Guthrie, 1988; Zhuang and Weiner, 1986]. Studies in the literature demonstrated that the binding to a non-specific 5' splice site could be possible [Hicks et al., 2010] as the interaction between the U1 snRNP and the 5' splice site is not highly complementary. For example; the recognition of this (5'ss) could occur when nucleotides that are not complementary bulge out of the interaction duplex resulting in better complementarity between the remaining nucleotides [Carmel et al., 2004; Roca et al., 2012]. Other evidence showed that the selection of the 5' splice

### 1.3 Splice Site Selection and Recognition

---

site through the U1-C subunit is also possible when there is not high complementarity between the U1 snRNAs and the 5' splice site [Zhu et al., 2001]. In addition, mutations in the 5' splice site and U1 snRNA in yeast showed that U1 snRNA is also essential [S eraphin et al., 1988; Siliciano and Guthrie, 1988]. The RNase H protection assay demonstrated the binding between U1 snRNP and the 5' splice site [G unzl et al., 2002]. In addition, the RNase H cleavage experiment showed that U1 snRNA is required for splicing [Kr amer et al., 1984].

Despite this, a highly suitable 5' splice site sequence that provides a high affinity for the U1 snRNP helps in selecting this 5' splice site. A splice site with such a consensus sequence is termed a strong splice site. On the other hand, sites with lower affinity sequences are termed weak splice sites. Accordingly, it was demonstrated that strong splice sites are favoured over weaker splice sites [Eperon et al., 1986; Lear et al., 1990], showing a poor relationship between the predicted U1 snRNP binding and 5' splice site usage. The usage of splice sites and their binding by U1 snRNPs does not always correlate [Eperon et al., 1993; Nelson and Green, 1990] and so it is not always the strongest site that gets used [Shapiro and Senapathy, 1987]. The relative position of potential 5' splice sites plays a role too. It has been demonstrated that the selection of a downstream 5' ss was always favoured when two strong 5' ss are occupied by U1 snRNPs. Such that, downstream splice sites are preferred even when a stronger competing 5' splice site inserted upstream can stimulate usage of the downstream splice site [Eperon et al., 1993; Hicks et al., 2010; Reed and Maniatis, 1986], showing a preference for intron definition over exon definition. Further research explained that the selection of 5' ss usage depends on the U1 snRNP binding rate at critical 5' ss during selection

### 1.3 Splice Site Selection and Recognition

---

[OMullane and Eperon, 1998]. Strong alternative 5' splice sites may be bound by U1 snRNPs concurrently and the selection of the 5' ss to be spliced depends on other factors. On the other hand, the use of weak 5' splice sites depends on the relative lifetimes of U1 snRNPs bound to them [Eperon et al., 1993, 2000; Hodson et al., 2012]. A single molecule study has demonstrated that both strong alternative (5'ss) were bound by U1 snRNPs in the early complex (E) and only one U1 snRNP remained at complex (A) [Hodson et al., 2012]. This indicates that recognition of the candidate 5' ss by U1 snRNPs is the first event and selection is a secondary event.

Other factors can occur and influence the splice site recognition and hence the splicing, for example the primary sequence of the RNA and its secondary structure [Buratti and Baralle, 2004]. In addition, recognition of the correct splice site can also be affected by other properties, such as intron length [Bell et al., 1998] or polymerase processivity [Nogués et al., 2003]. Another contribution can arise from the involvement of other sequences besides the ones at the 5' splice site.

Moreover, U6 snRNP is also important for the recognition of the 5' splice site. The interaction of U6 snRNP with the 5' splice site is enhanced by the presence of U1 snRNP, which in turn is displaced by U6 snRNP that binds the 5' splice site and stimulates the first transesterification reaction [Buratti et al., 2007; Grover et al., 1999].

The SR and the hnRNP proteins are also involved in the recognition of the 5' splice site if an alternative 5' splice site occurs. It was demonstrated that the SR proteins can bind in the proximity of the 5' splice site and can stimulate the use of the 5' splice site located downstream of their binding sites rather than the one upstream (if present) that will be inhibited (alternative

### 1.3 Splice Site Selection and Recognition

---

site) [Buratti and Baralle, 2004; Hicks et al., 2010].

A major study has been done on one of the major SR proteins, the SRSF1. The SRSF1 protein has an RNA recognition motif RRM domain, which interacts with the RRM domain of U1-70K (an auxiliary protein of U1 snRNP); this interaction is controlled by the phosphorylation of the SRSF1 RS domain. In conditions of hyperphosphorylation, the RS domain allows the RRM motif of the SR protein to bind the RRM motif of the U1 snRNP, when not phosphorylated the RS domain sequesters the RRM motif and there will be no binding [Cho et al., 2011; Krämer and Utans, 1991]. Interestingly, *in vitro* experiments over expressing SR proteins demonstrated that it is possible to stimulate the recognition of the 5' splice site even in the absence of U1 snRNP [Crispino et al., 1994; Tarn and Steitz, 1994]. In addition to the SR proteins, for instance, hnRNP factors are associated with two functions, the stimulation and inhibition of the selection of splice sites [Fisette et al., 2010]. The hnRNP A1 protein has been shown to modulate the selection of the 5' splice site. In particular, the hnRNPA1 can disrupt binding between U1 snRNP and the 5' splice site [Eperon et al., 2000; Pagani and Baralle, 2004]. In addition, it has been proposed that in some cases hnRNP A1 and hnRNP H promote the recognition of the 5' splice site. In these cases, hnRNP A1 and hnRNP H interact with their binding sites on intronic RNA, bringing distal 5' and 3' splice sites closer, favouring the splicing process [Fisette et al., 2010; Pagani and Baralle, 2004]. Finally, it has been shown that the presence of secondary structures close to the 5' splice can also affect its recognition. An example is seen in the tau gene where a mutation in the 5' splice site of exon 10 affects a stemloop structure [Zhang et al., 2005a; Zhang and Powell, 2005; Zhang et al., 2005b,c,d]. In some diseases that are associated with the presence of secondary

### 1.3 Splice Site Selection and Recognition

---

structure, such as ataxia telangiectasia mutated (ATM) and cystic fibrosis transmembrane regulator (CFTR), the deletion in the repressor element in ATM causes the use of an aberrant 5' splice site while, in CFTR, a new 5' splice site is created. In both cases, RNA secondary structure is the key regulatory element [Manley and Krainer, 2010].

#### 1.3.2 The 3' Splice Site

Compared with the 5' splice site (5'ss), there are three important and different conserved sequences involved in the 3' splice site (3'ss) selection: the branch point sequence (BPS), the polypyrimidine tract (PPT) and the AG dinucleotide at the (3'ss) [Reed, 1989], (Figure 1.8). The 3'ss and PPT for pre-mRNA splicing were identified in the 1980s [Mount, 1982; Reed, 1989; Shapiro and Senapathy, 1987]. The 3' splice site was recognized during the early step of spliceosome assembly by a heterodimer of two subunits of U2AF35 and U2AF65 [Gaur et al., 1995; Ruskin et al., 1988; Wu et al., 1999; Zamore and Green, 1989, 1991]. The subunits U2AF35 and U2AF65 are part of the auxiliary protein U2AF. Evidence shows that the mammalian branch point is specified primarily by its proximity to the intron/ exon junction, and the consensus sequence YNYURAC motif (R=Purine, Y=pyrimidine).

The branch point is recognized by the SF1 factor during early spliceosome assembly (the early complex (E)) [Berglund et al., 1997], which is displaced by the U2 snRNP in complex (A) (Figure 1.8) [Berglund et al., 1998; Krämer and Utans, 1991; Wahl et al., 2009]. Subsequently, recognition of the branch site involves binding of U2 snRNP through interaction with SF3a and SF3b to form the spliceosome complex (A) [Zhong et al., 2009].

On the other hand, the polypyrimidine tract sequence is rich in pyrim-

### 1.3 Splice Site Selection and Recognition

---

idines located between the branch point and the terminal AG at the intron/exon junction (the 3' splice site). It is required for splicing and is recognized by several proteins, such as polypyrimidine tract binding protein (PTB), the auxiliary factor U2AF65-kDa subunit and a splicing regulator Sex-lethal (Sxl) [Gooding et al., 1998; Green, 1991; Wagner and Garcia-Blanco, 2001]. Under normal splicing conditions, U2AF65 binds to the PPT during the formation of the early complex E [Kielkopf et al., 2004; Zamore et al., 1992]. Using purified U2AF65, UV cross-linking showed that U2AF65 can bind to the PPT in the absence of ATP at a wide range of temperatures and remains associated at the 3'ss [Zamore and Green, 1989]. The function of the U2AF65 is to bind the PPT and to bring the 3' splice site and the adenosine branch point closer to each other [Saulière et al., 2006]. In addition, it has been shown that deletion of the PPT prevents the formation of the lariat intermediate [Mullen et al., 1991; Roscigno et al., 1993]. The conserved terminal AG dinucleotide defines the 3' border of the intron. It plays an important role in splicing. This site is characterized by a short sequence (YAG/G) where Y denotes pyrimidine; the slash indicates the intron/exon boundary and the underlined nucleotides are conserved [Langford et al., 1984]. As the AG dinucleotides at the 3'ss are cleaved in the second step of splicing, it is also essential for spliceosome assembly at the earliest step. It has been demonstrated that the mutation of the AG dinucleotides at the 3' splice site can block the splicing reaction at the first step [Aebi et al., 1986; Lamond et al., 1987; Reed, 1989] when the PPT is strong.

## 1.4 Splicing Regulation

---

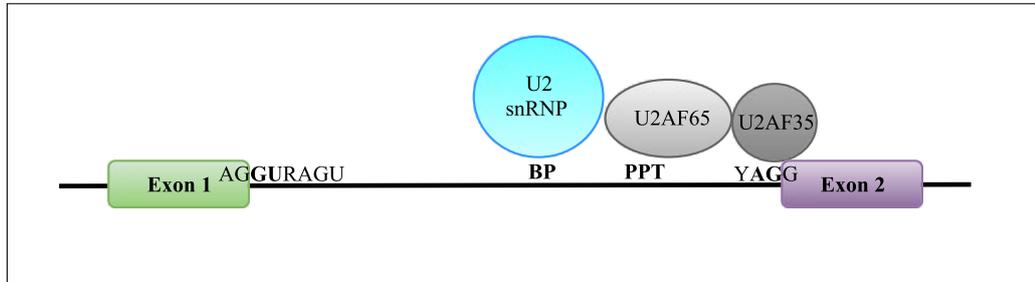


Figure 1.8: Schematic representations of the 3' splice site recognition. The figure represents the recognition of the 3' splice site by the U2 snRNP which binds the adenosine branch point (BP), by the U2AF65 which binds the polypyrimidine tract (PPT) and by the U2AF35 which binds the AG dinucleotide. The green and the purple boxes represent the exons. The solid black line represents the intron. The light blue circle is the U2 snRNP. The dark and light gray circles represent the U2AF complexes: U2AF35 and U2AF65, respectively.

## 1.4 Splicing Regulation

Apart from the splicing signals, further elements contribute to the definition of exon/ intron boundaries and the regulation of splicing. These splicing regulatory elements (SREs) in the case of cis-regulatory sequences are present in the pre-mRNA, which can vary regarding location and effect [Matera and Wang, 2014]. In general, SREs contribute to the recruitment of a set of proteins that can act as repressors or activators of splicing. These proteins termed as trans-acting splicing factors (SFs). They are acting typically by influencing spliceosome assembly.

A conspicuous example of the role of SREs is their contribution to the recognition of exon/ intron boundaries through a process termed exon definition [De Conti et al., 2013]. Such mode of splice site recognition is common in higher eukaryotes, where intron size exceeds that of exons. This mode could lead to splicing errors because of the existence of cryptic splice sites. Consequently, another class of SFs termed SR proteins (Serine-Rich

## 1.4 Splicing Regulation

---

proteins) promote the binding of snRNPs to the splice sites located at both ends of the same exon. This promotion occurs by binding to exonic splicing enhancers (ESEs). As a result, a cross-exon recognition complex is formed, which will ultimately lead to intron-spanning interactions through spliceosomal rearrangements. Conversely, intron definition emerges as the prevalent mode of splice site recognition in lower eukaryotes (short introns) [De Conti et al., 2013]. In this case, splice sites situated on both ends of the same intron are recognised without the help of SFs.

Together with the competition alongside splice sites and cis-acting SREs based on their sequence composition, the accessibility of those elements plays an important key role in alternative splicing regulation. Such accessibility can be influenced by the pre-mRNA secondary structure, chromatin arrangements and nucleosome positioning, and can also be dynamically controlled as a result of the co-ordination of the transcription and splicing processes [Brown et al., 2012; Plass and Eyraes, 2014].

In humans, a phenomenon is known as co-transcriptional splicing occur in most of the splicing events before transcription termination [Tilgner et al., 2012]. Such a phenomenon indicates that transcription elongation rates can have an influence on splice site choice: for example slow elongation will provide an opportunity for the recognition of weak splice sites while a fast elongation will promote the recognition of strong splice sites instead [Bentley, 2014]. Finally, the regulation of splicing is not limited to the role of specific SFs, because fluctuations in the concentration of core components of the spliceosome are also known to influence the splicing result [Saltzman et al., 2011]. Overall, these processes guarantee that splicing occurs in an accurate fashion. The accuracy of splicing is further increased by the many rearrangements that are required

## 1.5 The Additional Elements

---

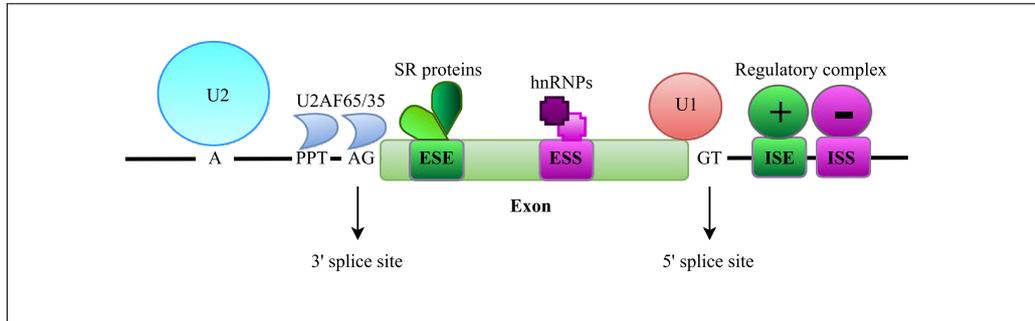


Figure 1.9: Schematic representation of the possible distribution of canonical and additional splicing cis-elements. The figure shows the elements involved in the splicing process. The canonical splicing signal that defines the exon boundaries are: the GT at the 5' splice site, the AG at the 3' splice site and the branch point adenosine.

before the actual intron removal can occur and splicing errors are rejected by the NMD pathway. On the other hand, the accumulation of splice site mutations or the alteration in the function of spliceosomal components can lead to phenotypic consequences and, in fact, dysregulation of splicing has been linked to many diseases, including cancer [Ladomery, 2013; Padgett, 2012; Tazi et al., 2009].

## 1.5 The Additional Elements

It was demonstrated that the splice site recognition occurs across the exons through the initial interaction between the 5' and 3' splice sites via the exon definition model [Berget, 1995]. This recognition is the result of a combinatorial regulatory mechanism [Smith and Valcárcel, 2000] that uses various controlling elements. In addition, other SREs can affect the splice site recognition too. They can act by increasing or decreasing exon recognition depending on their location, (Figure 1.9).

## 1.5 The Additional Elements

---

### 1.5.1 Regulatory Elements; Enhancers and Silencer

The 5' splice site, 3' splice site, branch point (BP) and polypyrimidine tract (PPT) are required and essential for splicing. Their sequences alter in nature, which could lead to unwanted or inefficient splicing or no splicing at all. In addition, it was found that the information within the 5' splice site sequence is not enough to detect them accurately. Further information is mandatory and has to contribute to determining them [Lim and Burge, 2001]. This information can be found in additional sequences identified as enhancers; these sites are named as (ESEs) when inside an exon, or intronic splicing enhancers (ISEs) when inside an intron. Opposite to the (ESEs), other factors can be present to suppress splicing. In that case, they are named silencers with the sequences termed (ESS) and (ISS) when inside an exon and an intron, respectively [Cartegni et al., 2002; Cartegni and Krainer, 2002].

#### 1.5.1.1 Splicing Enhancers

ESEs are recognised by different RNA binding proteins, which belong to the SR protein family of SFs. ESEs are usually located downstream of suboptimal 3' splice sites [Fairbrother and Chasin, 2000]. A purine-rich ESE has one or more binding sites for SR proteins, which are implicated in recruiting or strengthening the binding of SR proteins. For example, it was demonstrated that the ESE site recruits U2AF to the upstream 3' splice site and stimulates spliceosome assembly [Graveley et al., 1998; Graveley and Maniatis, 1998]. ESEs are involved in constitutive splicing events [Lavigne et al., 1993; Schaal and Maniatis, 1999a,b] and certain alternatively spliced exons [Lavigne et al., 1993; Ramchatesingh et al., 1995; Sun et al., 1993; Tian and Maniatis, 1993; Yeakley et al., 1996]. In addition, they were identified as reg-

## 1.5 The Additional Elements

---

ulators of alternative splicing [Black, 2003]. In addition, enhancer motif can affect the U1 snRNP and the selection of the 5' splice site through factors that bind to them. [Eperon et al., 1993] showed that SR proteins like SRSF1 can fulfil this role.

In previous studies, it was demonstrated that Tra2 and Tra2 of *Drosophila* bind to ESEs containing GAA repeats and, in conjugation with SR-family proteins, promote ESE-dependent splicing [Tacke et al., 1998]. Moreover, [Eldridge et al., 1999; Longman et al., 2001] showed that SRm160 protein is important for a typical ESE sequence that consists of GAA repeats to promote the splicing of a pre-mRNA derived from exons 3 and 4 of the *Drosophila* doublex gene [*dsx* (GAA)<sub>6</sub>]. Antibodies for SRm160 protein precipitated the (GAA)<sub>6</sub>-RNA but not the (GUU)<sub>6</sub>, indicating that SRm160 protein is specifically associated with GAA repeat sequences [Eldridge et al., 1999; Longman et al., 2001]. Increasing the GAA repeats from zero to six resulted in a significant stimulation of *dsx* pre-mRNA splicing in the presence of SRm160 protein [Blencowe et al., 1998; Eldridge et al., 1999]. In addition to interactions mediated through the ESE, the stable association of SRm160 protein with the *dsx* pre-mRNA could involve more interactions mediated through the formation of one or more snRNP-containing splicing complexes. For example in a mock-depleted reaction mixture, an antibody for SRm160 protein immunoprecipitated increasing levels of the *dsx* pre-mRNA as the numbers of GAA repeats in the ESE increased, whereas the *dsx* pre-mRNA with or without an ESE did not immunoprecipitate in the absence of U1 snRNP. In the absence of U2 snRNP, the level of *dsx* pre-mRNA immunoprecipitated increased as the number of GAA repeats increased. These results indicate that U1 snRNP and factors bound to the ESE are required to promote stable association of

## 1.5 The Additional Elements

---

SRm160 protein, whereas U2 snRNP is not. Furthermore, it was identified that the depletion of U1 snRNP prevented both the immunoprecipitation of *dsx* pre-mRNA in the presence of SRm160 protein and the binding of U2 snRNP to the *dsx* pre-mRNA even in the presence of GAA repeat. This indicates that both U1 snRNP and ESE components co-operate to recruit SRm160 and U2 snRNP to the *dsx* pre-mRNA. These results suggested a model in which SRm160 protein functions as a co-activator of ESE-dependent splicing by bridging between snRNP factors and SR proteins bound to ESE

Another example of an ESE is found in *Drosophila* doublesex (*dsx*) pre-mRNA. The ESE was found as a 13 nucleotide repeat element (*dsxRE*) and was essential for exon inclusion [Tian and Maniatis, 1993]. ESEs were demonstrated as domains enriched in purines and were initially found in immunoglobulin M. It was shown that the ESE is essential for the function of the upstream 3' splice site. In general, interruption or deletion of ESE sequences results in skipping of an exon during splicing.

In general, ESEs through SR protein binding create the process of splicing by recruiting SFs. The importance of ESE residues was investigated in this research in the presence and absence of the SRm160 protein and will be explained later in the chapter.

### 1.5.1.2 Splicing silencers

Splicing silencers are a group of regulatory proteins that repress or silence splicing. The sequence of these silencers can be purine or pyrimidine-rich and bind different SFs [Fairbrother and Chasin, 2000]. Several ways were identified in the regulation process via a silencer sequence. One of the ways is by antagonising the function of a nearby ESE. Alternatively, by recruiting fac-

## 1.5 The Additional Elements

---

tors that interfere with the splicing machinery by steric hindrance or through exon looping out of the pre-mRNA or by nucleation and co-operative binding [Cartegni et al., 2002; Cartegni and Krainer, 2002; Matlin et al., 2005]. Previous studies have identified that inhibition typically occurs during the initial ATP-independent recognition of splice sites. In addition, it has been demonstrated that the silencer motif sequences interact with the proteins belonging to the heterogeneous nuclear ribonucleoproteins family (hnRNP) [Cartegni et al., 2002; Cartegni and Krainer, 2002]. In addition, they can inhibit splicing. This family of proteins contains an RNA-binding domain and an auxiliary domain that is rich in amino acids that are responsible for protein-protein interactions [Dreyfuss et al., 2002; Smith and Valcárcel, 2000]. hnRNP proteins, similar to SR proteins, are responsible for various functions in the cell, including regulating transcription, pre-mRNA splicing and translation [Han et al., 2010]. One of the silencers, termed hnRNP A1, was shown to antagonise SRSF1 [Mayeda et al., 1993]. It is thought that the antagonising effect is through reciprocal blocking of binding sites [Eperon et al., 2000; Zhu and Krainer, 2000]. It also promotes exon skipping [Blanchette and Chabot, 1999].

Interestingly, it has been found that hnRNP A1 binds to both the 5' and 3' splice sites [Buvoli et al., 1990] and also binds to the U2 snRNP. This binding implies that hnRNP A1 is involved in the early stages of spliceosome assembly [Buvoli et al., 1992]. The classic model is that hnRNP A1 represses splicing by competition with SR proteins and promotes exon skipping by increasing the use of weak distal splice sites [Cáceres et al., 1994; Mayeda and Krainer, 1992; Mayeda et al., 1994].

[Eperon et al., 2000] identified that the splicing silencers can affect the selection of 5'ss. They showed that hnRNP A1 could shift the 5'ss selec-

## 1.5 The Additional Elements

---

tion from downstream to an upstream 5'ss by affecting the binding of the U1 snRNP. In addition, they could compete to bind to the same pre-mRNA with added hnRNP A1 and SRSF1 to represses or stimulate U1 snRNP binding at a downstream 5'ss. In addition, hnRNP A1 has also been found bound to an ESE at the 3' end of HIV-1 at exon 3 that could antagonise SR protein binding to an ESE at the 5' end of the exon and in turn inhibit the selection of an alternative 3'ss [Expert-Bezançon et al., 2004]. Moreover, the RNA secondary structure was found to be influenced by hnRNP A1 binding. hnRNP A1 could displace other bound proteins on RNA and spread along the RNA [Okunola and Krainer, 2009]. Another example of a splicing silencer is the hnRNP PTB; also named PTBP1 or hnRNP I [Cherny et al., 2010; Shen and Green, 2004; Shen et al., 2004]. PTB is a 57-kDa protein that contains four RRMs, which often bind to 3' splice site proximal regions. It binds to these polypyrimidine-rich sequences to repress splicing [Kafasla et al., 2012]. Different studies have reported PTB repressing splicing in several different ways. For example in some cases, PTB uses competitive inhibition, whereby it competes with U2AF65 [Saulière et al., 2006]. PTB has also been found to bind U1 snRNA, thereby possibly blocking the interaction between the U1 snRNP with 3' spliceosomal components [Sharma et al., 2011]. Moreover, it can also divert U2 snRNA to base pair to another sequence away from the branch point [Zheng et al., 2014].

## **1.6 Trans-acting Factors; Additional Proteins Involved in Splicing**

---

### **1.6 Trans-acting Factors; Additional Proteins Involved in Splicing**

The small nuclear ribonucleoproteins particles (snRNPs) and non-snRNPs are fundamental factors that are involved in the splicing process during spliceosome assembly. In addition, two other families of serine-arginine rich RNA-binding proteins have also been identified as the main components of regulatory complexes with functional specificity in the splicing process, and that usually bind the ESEs and ESSs [David and Manley, 2008].

#### **1.6.1 The Serine-Arginine Proteins; SR Family and SR-related Proteins**

A hallmark factor involved in RNA splicing is SR proteins, also named SR splicing factors (SRSFs) [Mayeda et al., 1999]. They are families of RNA-binding proteins, highly conserved in metazoan cells [Blencowe et al., 1999, 1998; Fu, 1995; Manley et al., 1996; Valcárcel et al., 1996; Valcárcel and Green, 1996] but apparently absent from lower eukaryotes such as *Saccharomyces cerevisiae* [Hutton et al., 1998]. Protein domains that are rich in serine-arginine dipeptide repeats facilitate protein-protein and protein-RNA interactions during spliceosome assembly, recruitment and catalysis [Cartegni et al., 2002; Cartegni and Krainer, 2002; Shen and Green, 2004; Shen et al., 2004; Valcárcel et al., 1996; Valcárcel and Green, 1996; Wu and Maniatis, 1993]. These proteins are known as SR or SR-family proteins. Another type of SR protein is also involved. They termed as SR related proteins.

SR proteins include SRp20, SRSF1/SF2/ASF, 9G8, SRp30c, SC35, HRS/SRp40, SRp55, SRP75 and p54. These proteins possess one or two

## **1.6 Trans-acting Factors; Additional Proteins Involved in Splicing**

N-terminal RNA recognition motifs (RRMs) allowing them to interact with pre-mRNAs in a sequence-specific manner [Buée et al., 2000] and a C-terminal domain rich in alternating serine and arginine dipeptides, the (RS) domain [Ars et al., 2000; Birney et al., 1993; Lin and Fu, 2006; Long and Cáceres, 2009]. Many of these serine residues are phosphorylated. Phosphorylation allows SR proteins to interact with each other. These protein-protein interactions are essential in spliceosome assembly and play an important role in promoting the recognition of splice sites [Amrein et al., 1994; Fetzter et al., 1997; Kohtz et al., 1994; Wu and Maniatis, 1993; Xiao and Manley, 1997, 1998]. It was also identified that SR protein activity is regulated by phosphorylation and dephosphorylation [Stamm, 2008]. For instance, it was demonstrated that in *in vitro* splicing, the RS domain phosphorylation is critical for spliceosome assembly and its subsequent dephosphorylation is important for catalysis to occur [Cao et al., 1997; Mermoud et al., 1992, 1994; Xiao and Manley, 1998]. The RS domain of these proteins could be phosphorylated and dephosphorylated via various proteins. They are named as kinases. Several kinases have been identified, including members of the SR protein kinase (SRPK), the CDC2-like kinase family (CLKs) and AKT family [Colwill et al., 1996; Duncan et al., 1997; Gui et al., 1994; Scott et al., 2012; Wang et al., 1998]. Variations in kinase levels or activity have been shown to influence splicing [Lin and Fu, 2006; Scott et al., 2012].

Members of the SR family and SR-related proteins were first identified through genetic analyses of *Drosophila*, and subsequently by their ability to influence splicing activity *in vitro* in HeLa extracts [Amrein et al., 1988; Ge and Manley, 1990; Goralski et al., 1989; Krainer et al., 1990]. Since their discovery, many additional members have been identified. SR and SR-related proteins

## **1.6 Trans-acting Factors; Additional Proteins Involved in Splicing**

represent a major key in constitutive and alternative splicing regulators [Lin and Fu, 2006; Long and Caceres, 2009] and function at multiple stages of spliceosome assembly from recognition of the 5' splice site at the earliest step of splicing to binding and regulation of exonic-enhancer sequences that stimulate the use of suboptimal splice sites by promoting interactions with each other and with snRNP-associated proteins containing RS domains [Graveley, 2000; Mueller and Hertel, 2011].

Moreover, SR-family proteins and other RS domains containing proteins also function in splice site recognition by interacting with specific exon sequences named enhancers (ESEs). They function by providing binding sites for regulatory factors to recruit association of parts of the splicing complex. Various lines of evidence suggest that SR proteins promote exon inclusion by binding to enhancer elements in exons. This effect on splicing is believed to happen through two mechanisms. First, through its RS domain, an ESE-bound SR protein can recruit or stabilize components bound to the adjacent 5' and 3' splice sites, such as U1 snRNP or the U2AF complex, and through direct interactions with the BPS and 5' splice site [Graveley, 2001; Graveley et al., 2001; Kohtz et al., 1994; Li and Blencowe, 1999; Wang and Manley, 1995; Wang et al., 1995; Zahler and Roth, 1995; Zuo and Maniatis, 1996; Zuo and Manley, 1994]. These interactions are important in pairing the splice sites across an exon through a process known as (exon definition) [Robberson et al., 1990]. Second, an ESE-bound SR protein can inhibit or block inhibitory SFs such as hnRNP proteins or other RNA-binding proteins [Kan and Green, 1999; Zhu et al., 2001].

Additionally, SR proteins function in ESE-dependent splicing in several ways [Blencowe, 2000; Graveley, 2000]. SR proteins can bind specific ESE

## **1.6 Trans-acting Factors; Additional Proteins Involved in Splicing**

sequences and recruit the splicing machinery via interactions of their RS domains with snRNP components such as U2AF35 and U1-70K [Graveley, 2001; Graveley et al., 2001; Lavigne et al., 1993; Wang and Manley, 1995; Wang et al., 1995; Wu and Maniatis, 1993; Zuo and Maniatis, 1996]. Some other SR-related proteins can function in ESE-dependent splicing by acting as splicing co-activators that bridge interactions between ESE-bound SR/ SR-related proteins and snRNPs such as SRm160/ SRm300 [Blencowe, 2000; Blencowe et al., 2000, 1998; Eldridge et al., 1999]. Binding of SR proteins can also enhance exon inclusion by antagonizing the activity of negative regulators bound at nearby silencer elements [Kan and Green, 1999]. Recent results also show that inclusion of an alternative exon can be repressed by strong interactions of SR proteins with the flanking constitutive exons [Han et al., 2011]. In addition to roles in alternative splicing regulation, some SR and SR-related proteins have been found to be involved in transcription, 3' end formation, mRNA export and translation [Blencowe et al., 1999; Long and Caceres, 2009]. SR proteins have also been found to affect the selection of 5'ss. Previous studies have shown that the intron-proximal 5'ss will be used when an ESE sequence is located between two alternative 5'ss and bound by SR proteins (SRSF1, SRSF2 or SRSF7) [Erkelenz et al., 2013; Spina et al., 2006; Wang et al., 2006]. In addition, SR proteins have been identified to interact with the U2AF component U2AF35 and the U1 snRNP component U1-70K. As the U1 snRNP binds at the 5' splice site and U2AF recognizes the polypyrimidine tract and 3' splice site, a model has been proposed that sees SR proteins bridging across the exon or intron [Wu and Maniatis, 1993].

Based on *in vitro* splicing assays, SR proteins appear also to be functionally redundant in their ability to accompany splicing in HeLa S100 extract

## **1.6 Trans-acting Factors; Additional Proteins Involved in Splicing**

[Fu et al., 1992; Mayeda et al., 1992]. In addition, SR proteins have also been found to play a role in other aspects of gene regulation, including translation [Sanford et al., 2004].

In addition to the SR proteins, many other SR-related proteins also function as regulators of splicing. These proteins contain RS and RRM domains, but with configurations different from the classical SR proteins. Examples of such SR-related proteins include Tra2 A and Tra2 B, which are homologs of transformer-2, an AS regulator involved in *Drosophila* sex determination. Other SR-related proteins contain RS domains alone or in combination with other RNA-binding domains [Blencowe et al., 1999]. SR-related proteins such as U2AF subunits, the U1 snRNP component U1-70K and the coactivators SRm160 and SRm300, may or may not have RNA-binding activity and are believed to function through protein-protein interactions across splice sites by binding to components of the splicing machinery and with other RS domain-containing proteins [Blencowe, 2000; Blencowe et al., 2000; Sharma et al., 2008].

### **1.6.2 SerineArginine-rich Splicing Co-activator SRm160 and Its Role**

To gain more insights into the structure and function of components involved in pre-mRNA splicing, numerous studies were done, ranging from using a bank of monoclonal antibodies, mAbs, for the recognition of proteins associated with splicing to immunoprecipitation assays and splicing. A candidate protein was discovered by [Blencowe et al., 1994], named SRm160 (SR nuclear matrix protein of 160-kDa), that consists of serine (S) and arginine (R) repeats units. Unlike the SR family, the predicted SRm160 lacks an RRM domain. The

## **1.6 Trans-acting Factors; Additional Proteins Involved in Splicing**

absence of this domain in SRm160 indicates that its association with splicing complexes is mediated by protein-protein interactions [Blencowe et al., 2000, 1998; Eldridge et al., 1999].

SRm160 functions as a co-activator of constitutive and exon enhancer-dependent splicing by bridging interactions between SFs that bind directly to pre-mRNA [Blencowe et al., 1999, 1998; Eldridge et al., 1999]. It is required for the splicing of specific pre-mRNAs [Blencowe et al., 1998]. It has been found that SRm160 is associated with different SR proteins and SR-family proteins that promote splicing by binding to different ESE sequences. Various SR proteins (SC35, SRp30c, SF2/ ASF and SRp20) were detected by immunoprecipitation assay via an antibody for SRm160 [Longman et al., 2001], which indicates that SRm160 promotes splicing activity by association with different SR and/ or SR-family proteins bound to ESE sequences. It was also found that an antibody to SRm160 coimmunoprecipitated from nuclear extract: 75-kDa and 40-kDa proteins which were referred to SRp75 and hTra2 $\beta$ , respectively. This immunoprecipitation was resistant to extensive pre-incubation of the nuclear extract with RNase, indicating that the association between htr2 $\beta$  and SRm160 is mediated by protein-protein interactions [Blencowe et al., 1999, 1998]. In previous studies, it was demonstrated that the hTra2 $\alpha$  and hTra2 $\beta$  of *Drosophila* Tra2 bind to ESEs containing GAA repeats and, in conjugation with SR-family proteins, promote ESE-dependent splicing [Tacke et al., 1998].

Moreover, it was identified that SRm160 is important for a typical ESE sequence that consists of GAA repeats to promote the splicing of a pre-mRNA derived from exons 3 and 4 of the *Drosophila* doublesex gene [*dsx* (GAA)<sub>6</sub>]. Antibodies for SRm160 precipitated the (GAA)<sub>6</sub>-RNA but not the (GUU)<sub>6</sub> indicating that SRm160 is specifically associated with GAA repeat sequences

## 1.6 Trans-acting Factors; Additional Proteins Involved in Splicing

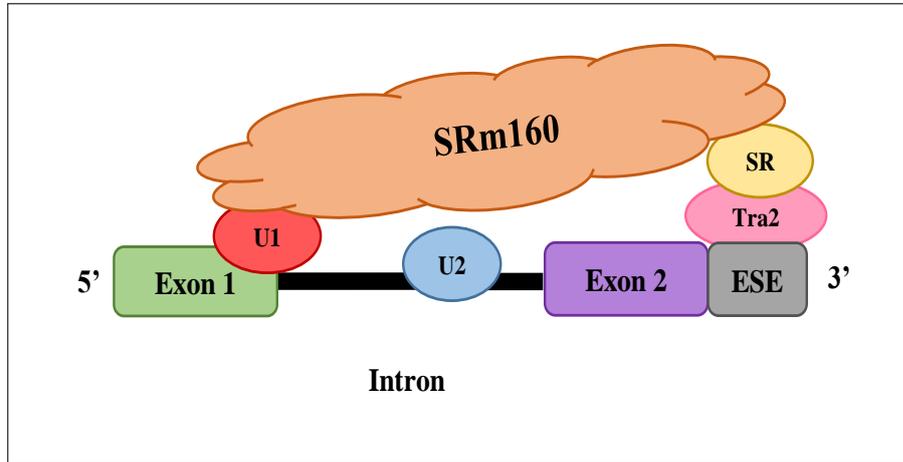


Figure 1.10: The role of SRm160 splicing co-activator in ESE function. It was proposed that SRm160 acts as a bridge factor through various interactions in order to promote the pairing of exons during the regulation of splice site selection. The binding of Tra2 to a GAA-repeat ESE and U1 snRNP to the 5' splice site is important to recruit SRm160 to pre-mRNA. These interactions simultaneously promote the binding of U2 snRNP to the pre-mRNA which also interact with SRm160.

[Eldridge et al., 1999; Longman et al., 2001]. Increasing of GAA repeats from zero to six resulted in a significant stimulation of *dsx* pre-mRNA splicing in the presence of SRm160 [Blencowe et al., 1998; Eldridge et al., 1999]. In addition to interactions mediated through the ESE, the stable association of SRm160 with the *dsx* pre-mRNA could involve more interactions mediated through the formation of one or more snRNP-containing splicing complexes. For example in a mock-depleted reaction mixture, an antibody for SRm160 immunoprecipitated increasing levels of the *dsx* pre-mRNA as the numbers of GAA repeats in the ESE increased. Whereas, the *dsx* pre-mRNA with or without an ESE did not immunoprecipitate in the absence of U1 snRNP. In the absence of U2 snRNP, the level of *dsx* pre-mRNA immunoprecipitated increased as the number of GAA repeats increased. These results indicate that U1 snRNP and factors bound to the ESE are required to promote stable association of

## **1.6 Trans-acting Factors; Additional Proteins Involved in Splicing**

SRm160, whereas U2 snRNP is not. Furthermore, it was identified that the depletion of U1 snRNP prevented both the immunoprecipitation of *dsx* pre-mRNA in the presence of SRm160 and the binding of U2 snRNP to the *dsx* pre-mRNA even in the presence of GAA repeat. This indicates that both U1 snRNP and ESE components co-operate to recruit SRm160 and U2 snRNP to the *dsx* pre-mRNA. These results suggested a model in which SRm160 functions as a co-activator of ESE-dependent splicing by bridging between snRNP factors and SR proteins bound to ESE (Figure 1.10) [Blencowe, 2000; Blencowe et al., 2000; Eldridge et al., 1999].

Another study demonstrated that depletion by RNA interference (RNAi) of SRm160 of *C. elegans* (CeSRm160), in combination with depletion one of the SR-family proteins such as CeSC35, caused a specific defect in the development of this organism, resulting in the production of unfertilized oocytes by the injected animal [Longman et al., 2001].

Interestingly, [Blencowe, 2000; Blencowe et al., 2000; McCracken et al., 2002; Szymczyna et al., 2003] demonstrated that SRm160 protein binds nucleic acids directly via the PWI motif within the highly conserved N-terminal domain. This motif facilitates the stimulation of 3' end formation by allowing the binding of SRm160 protein to transcripts. Moreover, the deletion of the N-terminal domain of SRm160 containing the PWI motif prevents SRm160 from stimulating 3' end cleavage *in vivo*.

The presence of the PWI motif in SRm160 suggests that both direct and non-specific binding of SRm160 to the pre-mRNA substrate can occur, disagreeing with the model stated in (Figure 1.11). Although the roles of its domains and the nature of its interaction are known, its binding pattern, however, remains unclear. There is nothing known about the actual number

## 1.7 Single-Molecule Studies

---

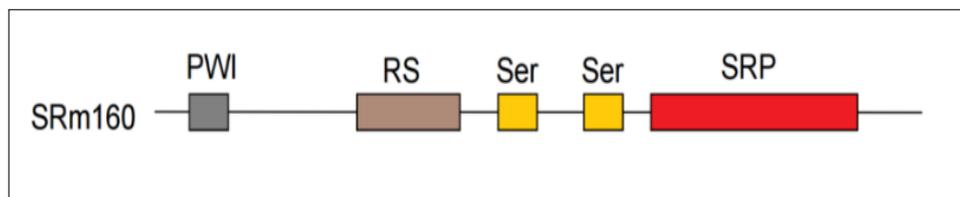


Figure 1.11: Schematic diagram shows SR domains in SRm160. This figure is adapted from [Blencowe et al., 2000]

of SRm160 molecules involved in splicing regulation and even less is known about the effects of different conditions.

A deeper understanding of SRm160 protein would greatly help. Single molecule methods provide the right tools to reveal the role played by SRm160 protein in splicing reaction.

## 1.7 Single-Molecule Studies

In 1926, Jean Baptiste Perrin was awarded the Nobel Prize in Physics. The study was on Brownian motion. This study confirmed the atomic nature of matter (from nobelprize.org).

Single molecule experiments are used to study and identify properties of individual molecules. For biological sciences, the single molecule approach meant a journey towards single molecule studies of the building blocks of life to understand nature at the smallest level. For a long time, the observation of single molecules was difficult, far out of reach and only ensembles of molecules could be investigated in bulk experiments. Those experiments allow an average measurement for the quantity observed, with statistical variations indicating the existence of values deviating from the measured mean. However, crucial information can be hidden within the averaged observations. The opposite of traditional bulk experiments is conceptually represented by single-molecule

## 1.7 Single-Molecule Studies

---

studies. Such studies describe a way to reveal the hidden information by distinctly interrogating each molecule, recording the result, exposing the full distribution and highlighting possible hidden heterogeneities. In the end, these individual measurements can be combined to produce the same averaged bulk results.

The last 25 years were the gate for providing a powerful perception beyond the measurement of averages. Single-molecule studies play an essential role in the fields of biology, chemistry and physics [Moerner, 2002]. Single molecule studies refer to a method involving optical detection of individual molecules, to study their kinetics, dynamics, assemblies and conformational rearrangements such as ribosomes, proteins and spliceosomes. The most widely used technique in single molecule research is molecular fluorescence, which is detailed later in this chapter.

Observing single molecules requires an approach that lets molecules emit photons themselves not only the reflection and absorption of light or electrons from molecules. Such process makes the detection of molecules independent of their size and also shifts the focus to sensitivity, for example detecting the luminescent signal. There are different types of luminescence used for microscopy [Créton and Jaffe, 2001; Melhuish, 1984]. Fluorescence is one of these types that "now permeates all of the cell and molecular biology" [Lichtman and Conchello, 2005]. This introduces spontaneous emission of light following the absorption of a photon by a molecule, first reported by [Stokes, 1852]. The photophysical pathways are listed in the following sections. They are also explained schematically in a simplified Jablonski diagram (Figure 1.12).

### 1.8 Molecular Fluorescence; Theoretical Background

In this section, the principles of fluorescence, colocalisation and fluorescence resonance energy transfer (FRET) are presented. First, photoluminescence is explained and the concept of electron energy levels in a molecule are introduced. The absorption and emission processes of a molecule are also described. Total internal reflection illumination method is outlined too. FRET and the factors that influence this effect are also described. Finally, an overview of some current applications of single-molecule microscopy for biological purposes such as RNA splicing is indicated.

#### 1.8.1 Absorption and Emission Processes

Luminescence (or photoluminescence) is the process in which a substance absorbs electromagnetic radiation of a certain energy and subsequently emits radiation of lower energy. Two types of photoluminescence have been identified. They are fluorescence and phosphorescence. In both cases, an electron is excited to a higher energy state and then returns to a lower energy state characterised by the emission of a photon. However, the processes by which the molecule emits the photon differ. The time between absorption and emission is extremely short, about  $10^{-8}$  s for fluorescence processes. In phosphorescent materials, molecules can be excited by photon absorption to energy levels, said to be metastable, which are states that last much longer, from  $10^{-3}$  s up to a few seconds or as long as minutes. In a collection of such molecules, many will de-excite to the lower state quickly, but many will remain in the excited state for much longer. Hence light will be emitted even after long periods. These

## 1.8 Molecular Fluorescence; Theoretical Background

---

substances are known more commonly as "glow in the dark" materials.

The energy levels in a molecule must be identified to understand the photoluminescence process. Molecules have multiple energy levels that depend on the structure of the molecule according to quantum theory. A molecule can absorb electromagnetic radiation by promoting an electron to a higher energy level (into an excited state). Multiple levels of energy have been classified; rotational, vibrational and electronic levels. Each of these transitions differs by an order of magnitude. Rotational transitions occur at lower energies  $10^{-3}$  eV (microwave region). Vibrational transitions occur in the infra-red  $10^{-1}$  eV. Electronic transitions occur in the ultraviolet and visible region of the electromagnetic spectrum, and these require higher energies (Figure 1.12).

Overall, the figure simplifies the photoluminescence pathway. A photon of suitable wavelength excites the molecule from the ground state into a higher excited singlet state. Internal rotational and vibrational relaxation lowers the energy of this excited state before the molecule returns to the ground state by immediately emitting a photon (fluorescence). Other energy transfer pathways are also possible [Lakowicz, 1999, 2013]. As the energy of the emitted photon is lower than the energy of the absorbed photon due to internal conversion and relaxations, the molecule releases the energy as a form of a photon and then returns to the ground state. The resulting wavelengths and the differences between them (the absorbed and emitted photons) are known as the Stokes shift (Figure 1.14). This shift leads to spectral separation of excitation and emission. Moreover, fluorescent energy is only released as photons of light. In a triplet state, the fluorophore can also decay through a non-fluorescent pathway.

## 1.8 Molecular Fluorescence; Theoretical Background

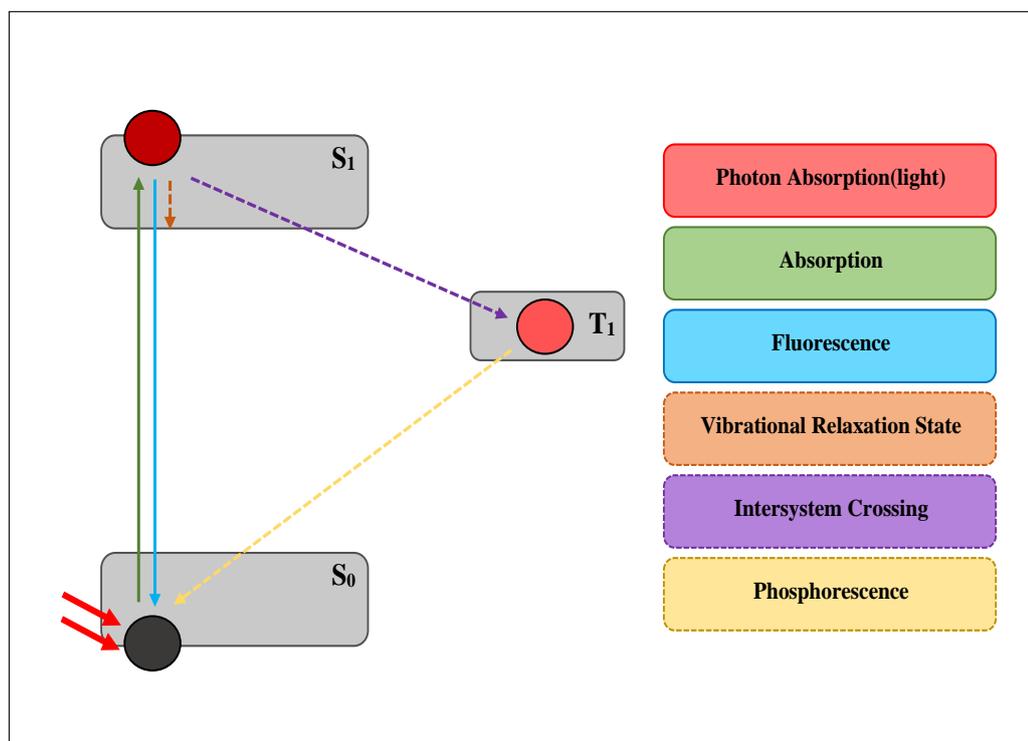


Figure 1.12: Simplified schematic diagram of Jablonski diagram. The molecule (black circle) at the ground state ( $S_0$ ). The molecule (red circle) at the first singlet state ( $S_1$ ) after irradiation by a light source (red arrow). To return to its ground state, the molecule emits energy by fluorescence or phosphorescence. Each coloured arrows relative to coloured boxes in the right side of the figure. Green arrow represents the transitions from ( $S_0$ ) to ( $S_1$ ); Absorption process. Blue arrow represents the returning to the ground state; fluorescence. Dashed orange arrow represents the internal mechanisms; vibrational and relaxation states. Dashed purple arrow represents the inter-system crossing mechanism. Dashed yellow arrow represents the phosphorescence process.

## 1.8 Molecular Fluorescence; Theoretical Background

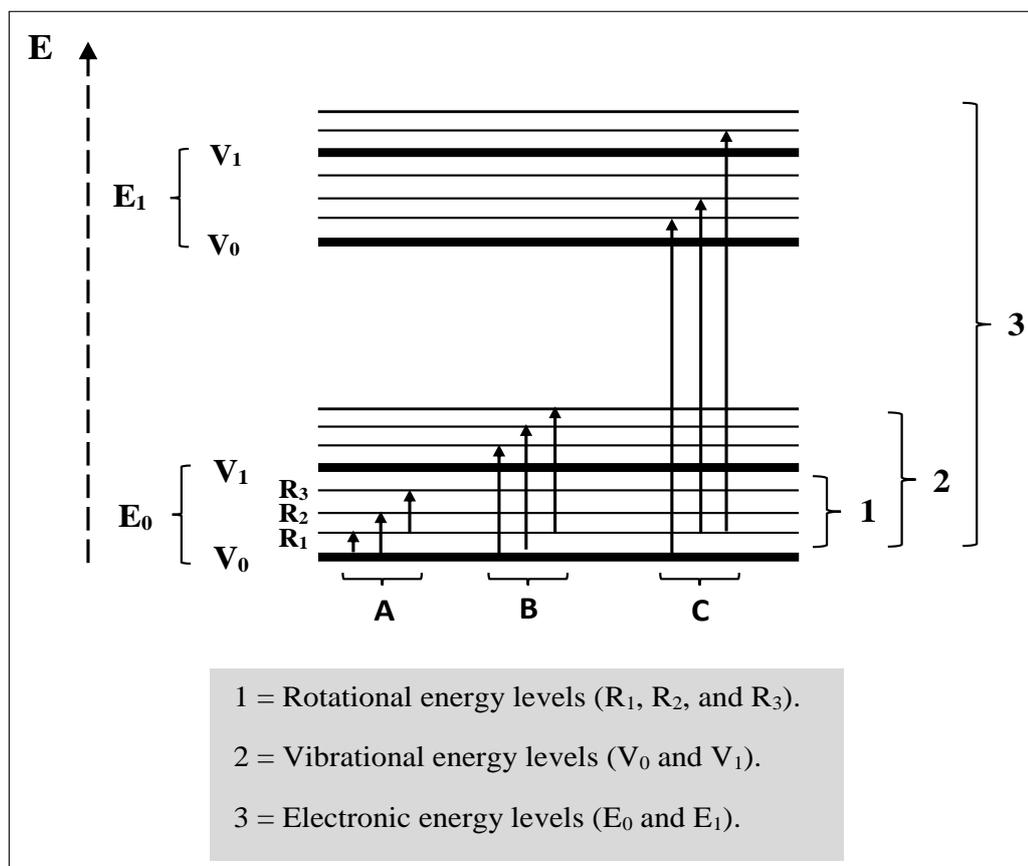


Figure 1.13: Simplified schematic diagram of transition energy levels. Dashed arrow represents the direction of the energy level ( $E$ ). ( $E_0$ ) and ( $E_1$ ) represent the electronic energy levels. ( $V_0$ ) and ( $V_1$ ) represent the vibrational energy levels. ( $R_1$ ), ( $R_2$ ) and ( $R_3$ ) indicate the rotational energy levels. A, B and C indicate the different options of molecule transitions.

## 1.8 Molecular Fluorescence; Theoretical Background

---

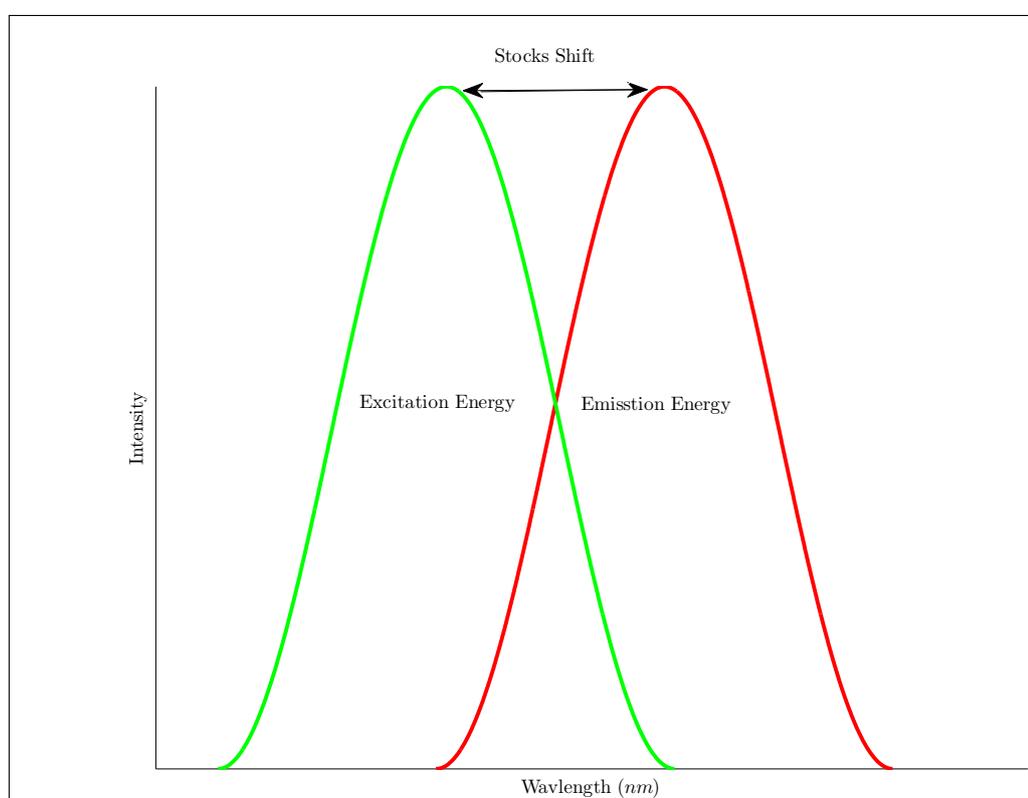


Figure 1.14: Schematic diagram of Stoke's shift. Green band represents the excitation spectra. Green band represents the excitation spectra. The difference in the wavelengths of the maxima for emission and excitation is the Stokes shift.

## 1.8 Molecular Fluorescence; Theoretical Background

---

### 1.8.1.1 Absorption Process

For fluorescence to occur, a molecule must first absorb a photon. If this photon has energy equal to, or greater than, the energy difference between the ground state and the excited state, it can be absorbed. Otherwise, the photon will be scattered. When a photon is absorbed, its energy is transferred to the valence electron and this electron is promoted to a higher molecular orbital, thus putting the molecule into an excited state. This transition is very fast, of the order of  $10^{-15}s$ .

The molar absorptivity ( $\epsilon$ ) of a substance is a measure of how strongly it absorbs light of a given wavelength and is usually measured in  $M^{-1}cm^{-1}$ . Transitions that have a high probability to occur have molar absorptivities of  $10^5 M^{-1}cm^{-1}$ , whereas for transitions that are theoretically forbidden, the molar absorptivity is less than  $10^{-4} M^{-1}cm^{-1}$  [Schenk, 1973]. It is possible to determine the molar absorptivity of a substance using a UVvis spectrophotometer. Such a spectrophotometer gives an experimental value of the absorbance using transmittance values. Transmittance is the ratio of the intensity of light transmitted through a substance  $I_1$ , to the intensity of light that initially fell on the surface,  $I_0$ , and can be explained by:

$$T = \frac{I_1}{I_0} \quad (1.1)$$

and absorbance (A) is defined as the negative logarithm of the transmittance,

$$A = -\log(T) \quad (1.2)$$

The Beer-Lambert law states that there is also a logarithmic dependence between transmittance and the product of both the absorption coefficient of the

## 1.8 Molecular Fluorescence; Theoretical Background

---

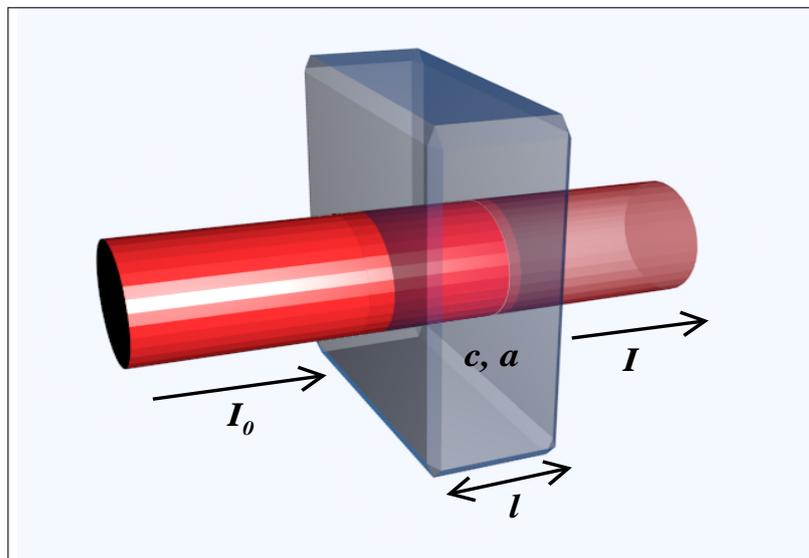


Figure 1.15: Schematic diagram of Beer-Lambert law. Gray cube represents the media (material). Red represents the laser (light) pathway. Black arrow represent the light direction.  $l$  and  $l_0$  represent and, respectively.  $l$  represents.  $c$  and  $a$  indicate the concentration of absorbing species in the material and the absorbance.

substance and the distance the light travels through the material,  $l$ , (Figure 1.15). The absorption coefficient is a product of the molar absorptivity,  $\epsilon$ , and the concentration of absorbing species in the material,  $c$ , [Rendell and Mowthorpe, 1987]. This gives the relationship below,

$$T = 10^{-a \cdot I} = 10^{-\epsilon \cdot c \cdot I} \quad (1.3)$$

Combining Equation 1.2 and Equation 1.3 gives

$$A = \epsilon \cdot c \cdot I \quad (1.4)$$

## 1.8 Molecular Fluorescence; Theoretical Background

---

### 1.8.2 Emission Process

A Jablonski diagram (Figure 1.12 and Figure 1.13 ) illustrates the electronic states of a molecule and the transitions between them. The states are arranged vertically by energy. Electronic energy levels with the same spin as the ground state are called singlet states and are indicated by the letter S [Schenk, 1973]. Electronic energy levels with a different spin from the ground state are called triplet states and are indicated by the letter T. Photo-emission processes are typically dominated by transitions between singlet and triplet states. Non-radiative transitions are indicated by sinusoidal arrows and radiative transitions by straight arrows. The vibrational ground states of each electronic state are indicated with thick lines, the higher vibrational states with thinner lines. The absorption and de-excitation processes of the molecule, as described above, are shown in the Jablonski diagram in (Figure 1.12 and Figure 1.13 ).

When an electron in a molecule has been energetically promoted to the excited state through the absorption of radiation, it returns to the ground state through radiative and non-radiative pathways. The radiative pathways involve photon emission and non-radiative pathways include energy transfer through collisions, resonance energy transfer through near field dipole-dipole interactions, photochemical decomposition, etc. A change in the vibrational and rotational states of the molecule can also cause a loss of energy via a non-radiative route. In the case of fluorescence the electron stays in the  $S_1$  excited state typically for  $10^{-8}s$  and returns to the ground state  $S_0$  very quickly, on the order of  $10^{-15}s$ . In phosphorescence, the molecule undergoes inter-system crossing, where the electron in the  $S_1$  state changes its spin and therefore its energy and relaxes into a triplet state  $T_1$ , this is illustrated in Figure 1.12. De-

## 1.8 Molecular Fluorescence; Theoretical Background

---

excitation from this triplet state to the ground state results in the emission of a photon. The emitted photon will have less energy than the photon that was absorbed due to rotational and vibrational changes. The absorption and emission of energy are unique characteristics of a particular molecular structure. The difference in energy (or wavelength) between the absorbed photon and the emitted photon is known as the Stokes shift, as shown in (Figure 1.14). A large Stokes shift is often highly desirable as it reduces the need for optical filters, which are used to separate exciting light and fluorescence emission.

### 1.8.3 Fluorophores

Some biological structures show auto-fluorescence, such as mitochondria and lysosomes that contain endogenous fluorophores like pyridinic and flavin coenzymes [Monici, 2005]. Tryptophan, tyrosine and phenylalanine (aromatic amino acids) show fluorescence in the near ultraviolet too [Teale and Weber, 1957]. These are exceptions, and most biological molecules do not fluoresce by themselves. Therefore, a fluorescent molecule, termed fluorophore, has to be attached to them. There are numerous strongly fluorescent proteins that cover a wide range of wavelengths [Shaner et al., 2005], such as the green fluorescent protein (GFP) [Day and Davidson, 2009] and mCherry [Graewe et al., 2009]. Alternatively, artificial fluorescent dyes, which provide superior photostabilities, can be also used to label a biological molecule, such as the Cyanine [Mujumdar et al., 1993] ATTO dyes [Buschmann et al., 2003] and Alexa Fluor [Berlier et al., 2003]. Recently, quantum dots have expanded the possibilities for labels even further [Resch-Genger et al., 2008].

These fluorophores usually have to be attached to the molecule of interest. This can be done in multiple ways. Immunofluorescent staining using

## 1.8 Molecular Fluorescence; Theoretical Background

---

fluorescent antibodies [Miller and Shakes, 1995] or with fluorophore-specific labelling [Haugland, 2002; Wiederschain, 2011] are examples of these ways. In the case of fluorescent proteins, new combined proteins have been created that contain the fluorescent protein directly attached to the protein of interest. In such method, the label is expressed by the cell alongside the protein [Chalfie, 1994]. To label DNA or RNA, modified single-stranded nucleotides with attached fluorophores can easily be annealed. A hybrid 2'-O-Me/ locked nucleic acid oligo is a good example that has been used for labelling [Crawford et al., 2008].

Despite this, problems of fluorophore attachment can exist. For example the complex formed between a fluorescent label and protein has different properties from the protein alone, at least differing in molecular weight. These differences might affect the folding and function of the protein, which can lead to observed different behaviours. Examples include false colocalisation in *Escherichia coli* cells [Landgraf et al., 2012], changes in association profiles of glycan-binding proteins [Fei et al., 2011] and an altered charge of bovine serum albumin [Bingaman et al., 2003]. It is, therefore, essential to investigate whether the results are independent of the fluorescent label or not.

Despite that, fluorescence microscopy is a popular technique, and multifaceted applications exist that use fluorescence for single molecule experiments. Some of these are described below.

Forster resonance energy transfer (FRET) is one of the single molecule techniques. It is a method to measure distances below 10 nm that takes advantage of dipole-dipole interactions between two fluorophores. In this method; the energy of one excited donor fluorophore is transferred to a second acceptor fluorophore. The efficiency of this transfer is heavily dependent on the

## 1.9 Total Internal Reflection Fluorescence Microscopy

---

distance between those two fluorophores and limited to small distances below 10 nm. It enables intramolecular distances to be measured or site-specific binding between molecules to be detected. It is a general technique that has been adapted to single molecules [Roy et al., 2008]. Single molecule FRET (sm-FRET) can be used to obtain kinetic constants for protein folding [Schuler and Eaton, 2008] or DNA unwinding by helicases [Myong et al., 2007]. One of the advantages of using FRET is the ability to investigate the three-dimensional structural dynamics of proteins [Forkey et al., 2003]. In addition, the time between absorption and emission of a photon contains further useful information about internal properties such as rotation and the likelihood of inter-system crossings and external features like the formation of excimers [Berezin and Achilefu, 2010]. Another method termed fast alternating laser excitation (ALEX) spectroscopy observes the emission of donor and fluorescence concurrently [Kapanidis et al., 2005]. Such a process allows the identification of the FRET efficiency and donor/ acceptor relative stoichiometry by measuring the acceptor emission directly by laser excitation and indirectly by FRET.

## 1.9 Total Internal Reflection Fluorescence Microscopy

Total internal reflection fluorescence (TIRF) microscopy applies an evanescent wave to illuminate a region of a sample placed on a surface. When the laser beam (exciting laser) hits the interface between two media, where the refractive index decreases on the other side (the second media) with an angle above the critical angle, it is reflected back to the first medium [Karlström and Nygren, 2001]. When this occurs, an evanescent electric field is generated that

## 1.9 Total Internal Reflection Fluorescence Microscopy

---

penetrates into the second medium by (50-150 nm). The exciting laser beam has refracted according to Snell's law,

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2) \quad (1.5)$$

where  $\theta_1$  and  $\theta_2$  are the angles of incidence and refraction, respectively. The  $n_1$  and  $n_2$  represent the corresponding refractive indices of the media. For a laser beam passing through a medium with a higher refractive index such as glass, into a medium with a lower refractive index, like water  $n_1 > n_2$ , the angle of refraction is larger than the angle of incidence  $\theta_1 > \theta_2$ . For large enough angles, the light is completely reflected at the interface (Figure 1.16).

The minimum angle at which this total internal reflection occurs is termed the critical angle. Above this angle, the incident light is totally reflected.

$$\theta_C = \arcsin\left(\frac{n_2}{n_1}\right) \quad (1.6)$$

An important feature of total internal reflection is the generation of an evanescent wave that propagates into the lower refractive index medium. In biological microscopy, such medium would be an aqueous sample. The advantage of this illumination is the extremely restricted illumination of the sample, reducing the possible background from the rest of the sample. Moreover, multiple ways exist to utilise total internal reflection as a method of fluorescence microscopy. One is to apply an alternative high refractive index object, commonly a prism [Axelrod, 2001; Thompson and Steele, 2007] or the microscope objective directly [Thompson and Steele, 2007]. In the objective-based method, the laser beam is passed through an objective lens, which needs to have a numerical aperture (NA) of more than 1.4 to achieve TIRF. The laser

## 1.9 Total Internal Reflection Fluorescence Microscopy

---

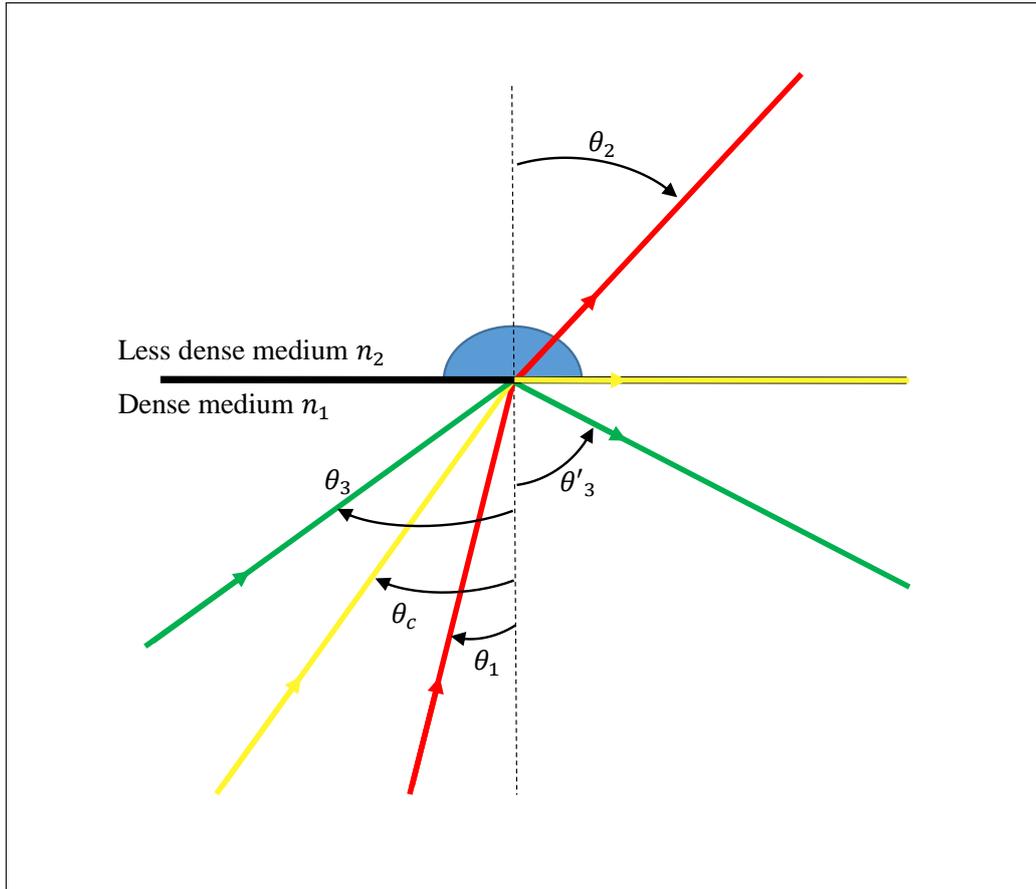


Figure 1.16: Schematic diagram of refraction at the interface of two different media. Medium with higher refractive index is indicated as ( $n_1$ ) below the solid black line. Medium with lower refractive index is indicated as ( $n_2$ ) above the solid black line. Light path shown in red is for normal reflection with an incident angle smaller than the critical angle  $\theta_1 < \theta_c$ . Yellow path shows the reflection parallel to the boundary for an incident angle equal to the critical angle. Path shown in green is for an incident angle greater than the critical angle  $\theta_3 > \theta_c$ .

## 1.9 Total Internal Reflection Fluorescence Microscopy

---

beam is passed through an immersion oil and a glass cover slip to hit the solution of a higher refractive index at an angle above the critical angle, producing the evanescent field through TIRF. The beam returns through the same objective lens. This method has some limitations, including a low signal-to-noise ratio due to fluorescence from the immersion oil and reflections from the laser beam that is passed through the objective [Fialcowitz et al., 2005; Moerner and Fromm, 2003]. On the other hand, the prism-based method differs from the objective-based TIRF method through the fact that the laser is passed hitting a prism. The general idea involves propagating laser light through a prism containing a liquid immersion oil that corresponds to the refractive index of the glass cover slip ( $n = 1.51$ ). The fluorescence then is collected through the objective. The detector of choice for wide-field microscopy is a CCD (charge coupled device). Images are produced by integration on a 1-1000 ms timescale. As there are entirely separate beam paths for excitation and emission and a larger possible illumination area, the prism-based illumination has the advantage of lower background. In contrast, the objective-based illumination makes use of objectives with a higher magnification and a higher NA, collecting more light and achieving a higher collection efficiency per single molecule [Hassler et al., 2005]. Usually, the sample is more accessible with objective illumination, allowing uncomplicated manipulations.

Because of the advantages stated above, we decided to choose objective-based illumination (Figure 1.17). The objective is positioned below the sample interface and coupled to the cover slip with immersion oil, all three having comparable refractive indices. The incident laser beam is guided into the objective parallel to the optical axis, focused onto the back focal plane. Moving it outwards results in different incident angles of collimated laser light at the

## 1.9 Total Internal Reflection Fluorescence Microscopy

---

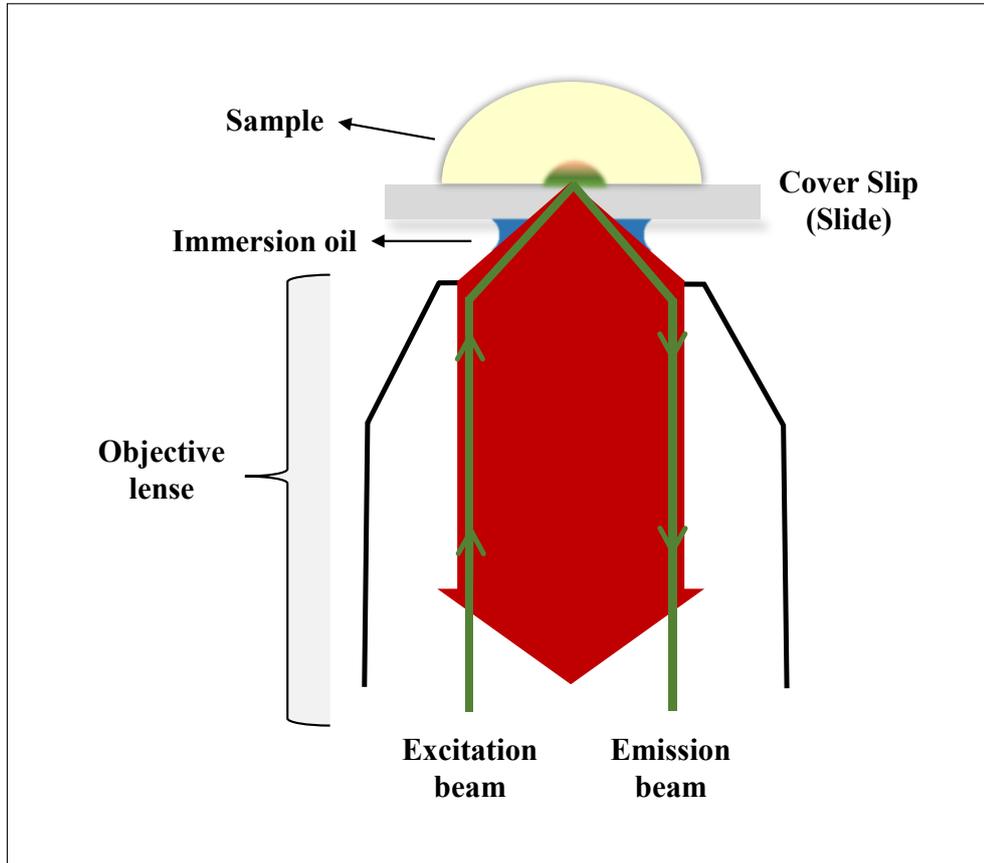


Figure 1.17: Schematic diagram of the objective-based illumination microscopy. Excitation and emission beam pathways are shown in green and red, respectively. All parts are labelled.

cover slip sample interface. The incident beam is reflected at the interface, giving rise to the aforementioned evanescent wave penetrating into the sample.

Overall and with its selective surface-defined illumination, TIRF is especially well suited to investigate membranes [Thomson et al., 1993] and membrane-related events like cell signalling [Sako et al., 2000]. For example the movement of single molecules that have been confined to the membrane has been tracked [Schmidt et al., 1996]. In addition, individual steps of myosin V moving along actin filaments can be seen [Yildiz et al., 2003]. Another implementation of it is for colocalisation experiments, which are the primary

## 1.10 Colocalisation

---

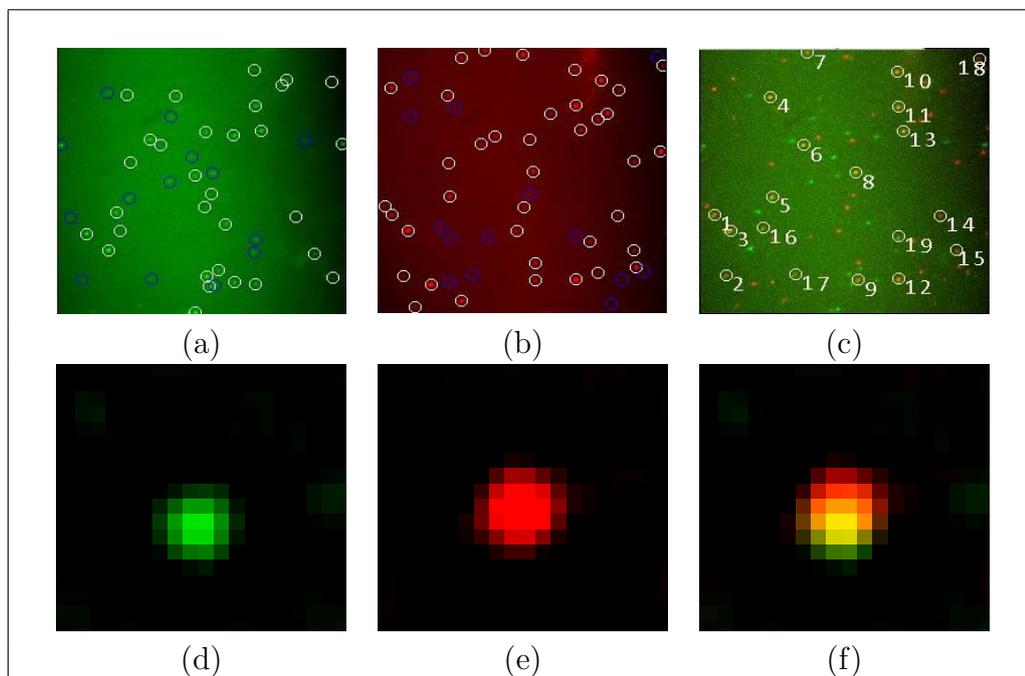


Figure 1.18: Simplified diagram of colocalisation. From (a) to (c) are Lab-View overview which represents green spots (GFP at 488nm), red spots (Cy5 at 633nm) and the colocalised spots (overlapped spots). (d) indicates the exemplary superposed image of a green spot. (e) indicates the exemplary superposed image of a red spot. (f) indicates the superposed image (d) and (e) that shows the colocalized spots in yellow.

research in this thesis and are discussed below.

## 1.10 Colocalisation

In a single molecule study or even in fluorescence microscopy, colocalisation refers to the observation of the spatial overlap between spectral fluorophores. It also describes the presence of two or more fluorophores attached to the same molecule. In the context of fluorescence imaging, it means that the fluorescent signals emitted by the fluorescent molecules occupy the same pixel in the image (in close proximity to each other), where the fluorescence emissions from their attached fluorophores overlap, see figure (1.18).

## 1.10 Colocalisation

---

To illuminate molecules for colocalisation studies, confocal illumination can be used for detecting dual-labelled complexes in solution. Such illumination can be performed using ALEX [Kapanidis et al., 2005] or by using ultra-fast laser pulse excitation [Olofsson and Margeat, 2013]. Surface selective illumination is preferred, where the complexes are immobilised on the surface. Illumination can be achieved by total internal reflection. The work presented later in this research has mainly done with the surface illumination by total internal reflection.

Moreover, to assess colocalisation, first, the spots that indicate a labelled molecule have to be identified, the position of each fluorophore has to be obtained, and then the distance between colocalised molecules can be calculated. All this and further analysis can be achieved via algorithms and computer analysis.

### 1.10.1 Photobleaching

Of concern from colocalisation measurements, the percentage of colocalised molecules is essential information but not only that can be extracted from the data. The intensity of fluorescence molecules over time contains additional information about the number of colocalised fluorophores.

Each fluorophore has an individual quantum yield. The quantum yield indicates the ratio of decay occurring by fluorescence and non-fluorescent pathways. Different quantum yields result in different numbers of emitted photons for the same laser intensities [Weber and Teale, 1957]. The detection efficiency, which indicates how many of the emitted photons are collected by the set up, is of concern from the experimental side. If the presence of an additional fluorophore results in a big change in signal level, it becomes possible to count the

## 1.10 Colocalisation

---

fluorophores. If multiple fluorophores suddenly stop emitting photons, one at a time, simply counting the successive downward steps in the intensity-time trace, the trace would tell us the number of fluorophores present initially. Almost all fluorophores stop emitting photons at some point due to a process termed as photobleaching [Dempsey et al., 2009; Widengren and Rigler, 1996]. This causes permanent loss of their ability to fluoresce. The average time until bleaching depends on the excitation intensities and other environmental factors, e.g. oxygen concentration, temperature and pH [Dempsey et al., 2009]. If the average time until bleaching is adjusted by either controlling the laser intensity or the environmental conditions, the separate bleaching events can be distinguished [Casanova et al., 2007; Leake et al., 2006]. Each bleaching step corresponds to one fluorophore, and each fluorophore is bound to one molecule, so each bleaching step indicates that one labelled molecule is present.

Colocalisation studies on single molecules have been used widely to look at a wide range of biochemical processes, such as transcription initiation [Friedman and Gelles, 2012], ribosomal translation of mRNAs [Tsai et al., 2012] and spliceosome assembly [Hoskins et al., 2011b]. Later in this research, single molecule studies alongside colocalisation were applied to study the behaviour of a coactivator SR splicing factor in RNA splicing (SRm160).

To assess colocalisation and extract information of interest from raw data, single molecule experiments have been developed, accompanied and supported by computer-based algorithms. This development was kindly performed by Robert Weinmeister [Weinmeister, 2014] and was applied in this research thesis to detect and exclude information of every single molecule of interest. Further development was applied to the program with thanks to Dr. Alaa Kahdidis to reduce substantially the time needed to analyse the data.

### 1.11 Fluorescence Resonance Energy Transfer

The use of two fluorescing dyes (fluorophores) makes it possible to identify the structural features of molecules via a mechanism termed FRET.

FRET is a mechanism describing energy transfer between a fluorescent molecule (donor; D) in an excited electronic state to a second molecule (acceptor; A) [Förster, 1955; Lakowicz, 2013]. The donor molecules emit at shorter wavelengths, which overlap with the absorption spectrum of the acceptor. The energy is not emitted by the donor (D) as a photon, or absorbed as a photon by the acceptor (A), but is transferred by non-radiative pathways. The perturbation by the excited donor (D) molecule and the acceptor (A) molecule takes place electrostatically as a dipole-dipole interaction. The emission band of the donor (D) must overlap sufficiently with the absorption band of the acceptor (A) for FRET to occur (Figure 1.19). The rate of such transfer of energy is dependent on the separation of the two molecules. This interaction only takes place in the range (10-100 Å), limiting the distance over which FRET can occur. The physics and chemistry of FRET have been well studied theoretically for years [Kubin and Fletcher, 1983; Lakowicz, 2013] but only with recent technical advances has it become achievable to apply the FRET mechanism to biomedical research [Shera et al., 1990].

FRET efficiency is heavily dependent on the inverse sixth power of intermolecular separation ( $1/(\text{distance})^6$ ), which makes it sensitive for investigating a variety of biological phenomena that give rise to distance changes of this order. Therefore, one of the main applications of FRET is as a spectroscopic ruler, which can probe distances on the nanometre scale through fluorescence measurements, based on Förster's basic rate equation for a donor and acceptor

## 1.11 Fluorescence Resonance Energy Transfer

---

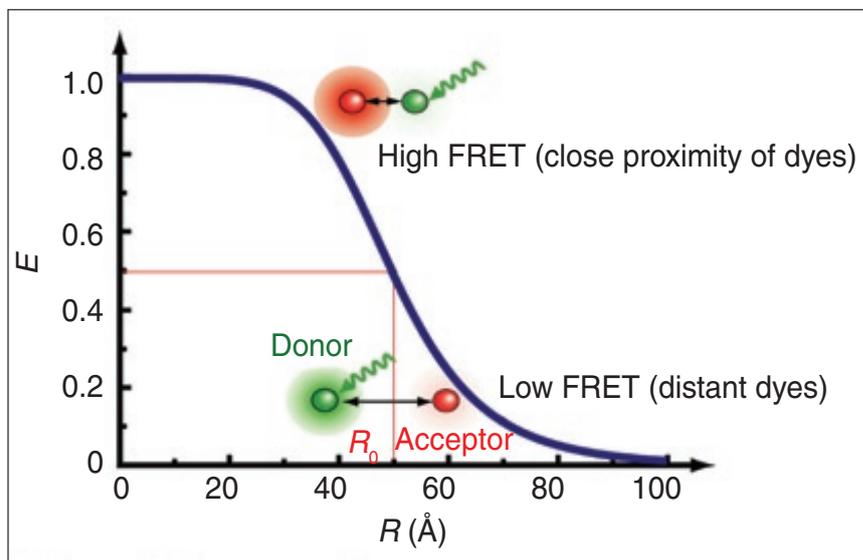


Figure 1.19: Schematic diagram of the mechanism of FRET.

at a distance ( $r$ ) from each other.

$$E = \frac{R_0^6}{R_0^6 + R^6} \quad (1.7)$$

As with all biological applications, sensitivity is essential. There are many ways of increasing the FRET efficiency. The FRET efficiency has been identified by the quantum yield of the energy transfer transition. It can be increased with an increase in the percentage of donor and acceptor overlap spectrum, by using different and suitable dye pairs. However, this also causes an increase in the bleed-through signal in the FRET channel. The efficiency of FRET could also be improved by optimising the concentration of the fluorophore used. While highly efficient traditional FRET systems are beneficial, it would be desirable to have a FRET system that was not limited to the 10 nm interaction distance. This would enable molecular interactions to be investigated that occur beyond this distance. Some work has been done to achieve this goal, but so far nothing greater than a 75% increase in the Forster radius

## 1.12 Single Molecule Experiments on RNA Splicing

---

has been reported.

Suitable donor/ acceptor pairs require sufficient spectral overlap of the emission and absorption spectra, high quantum yield of a donor, high absorption coefficient of acceptor and matching orientations of the donor and acceptor dipoles. Generally, for FRET to occur, the acceptor does not have to be fluorescent, but if it is, it emits a photon following transfer, which is characteristic of the acceptor fluorescence spectrum, although only D has been initially excited. Additionally, if the distance between the donor and acceptor molecule is beyond the FRET limit and the system is excited, donor emission can be observed but no energy transfer occurs and there is no emission of the acceptor. However, when the donor and acceptor are within the FRET limit, energy transfer can occur and both donor and acceptor emission can now be seen, with the donor emission reduced.

## 1.12 Single Molecule Experiments on RNA Splicing

The first observation of RNA splicing at single molecule level in real time was done by [Crawford et al., 2008] using a TIRF microscope. This experiment was done in whole cell extracts of yeast. They used RNA substrate with different fluorophores (dyes) that were incorporated into its exons and introns. RNA was attached to a surface glass through PEG-biotin and streptavidin. By observing a number of single molecules of RNA and analysing colocalisation over time, a loss of intron-specific fluorescence was shown, indicating the progress of splicing of the fluorescently labelled RNA.

The first insights into the dynamics of the yeast spliceosome itself were

## 1.12 Single Molecule Experiments on RNA Splicing

---

revealed by [Hoskins et al., 2011a,b]. They observed its sub-complex dynamics in real time in yeast whole cell extract. RNA substrates were bound to a PEG-modified glass surface and illuminated using TIRF with three different wavelengths. They observed colocalisation of pre-mRNA with splicing components, such as U1 snRNP, which was monitored over time. To distinguish between bleaching and dissociation of the components, the splicing components were labelled with two fluorophores of the same kind. If the observed intensity dropped by half, it indicated a bleaching event. If the intensity reduced to zero, either both fluorophores had bleached at the same time or, more likely, the component had dissociated. This allowed the authors to follow association and dissociation events over time and calculate reaction constants. They found an association of U1 snRNP only in the presence of a 5' splice site. Only in the presence of (ATP) did the sub-complexes U2 snRNP, tri-snRNP and NTC associate, and each showed different rates. They found this to be consistent with an assembly order of U1 U2 tri-snRNP NTC. All associations were reversible and only increased the splicing committance, no step guaranteed a proceeding splicing event.

In addition, a TIRF microscope was used to identify the number of bound proteins to observe the stoichiometry of these proteins involved in mammalian splicing. This process involved counting bleaching steps and analysing the total emission intensity. [Cherny et al., 2010] were the first to assess it by counting the number of PTBs bound to the pre-mRNA in nuclear extract. Similarly, the stoichiometry of U1 snRNPs in the early splicing complexes and their role in splice site selection was observed [Hodson et al., 2012]. They used different substrates with either strong or weak 5' splice sites. They identified that for two possible 5' splice sites there can be two bound U1 snRNPs in the

## 1.12 Single Molecule Experiments on RNA Splicing

---

early complex (E) with only one U1 snRNP remaining in complex (A).

A different approach in the TIRF is single molecule FRET studies. These studies can provide information about distances during splicing and conformational changes. The previous study by [Hoskins et al., 2011b] was subsequently extended to include FRET [Crawford et al., 2013]. This combination in their study allowed the authors to follow the spliceosome assembly over time and the conformational changes that occurred during the same time. The FRET measurements revealed that the 5' splice site and branch site remain separate during the spliceosome assembly and only approach each other just before the first catalytic step, after which the process becomes highly dynamic. This was contrary to previous bulk studies that suggested that this conformational change occurs around complexes E or A. In addition, [Abelson et al., 2010] observed time-dependent and (ATP)-dependent conformational transitions, all of them reversible with a highly fluctuating spliceosome operating. They labelled the RNA at the 3'- and 5'- exons with fluorophores and RNA was attached to the surface through the 3'-exon. [Lamichhane et al., 2010] also used single molecule FRET to investigate the activity of pure PTB on short RNA oligonucleotides.

All of these single molecule studies on RNA splicing have the same common methodology. Briefly, to follow the same molecule over time, the pre-mRNA is immobilized on a surface, the sample is illuminated using total internal reflection illumination. The analysis is done by colocalisation studies, counting bleaching steps and further analysis by looking at intensity-time traces.

### 1.13 Contributions of Thesis

This thesis proposes novel observation for a co-activator protein (SRm160) using single molecule techniques, such as, single molecule colocalisation and FRET methods. The advantages of using single molecule approaches are: (1) they are the only way to get data on stoichiometry, for example, these methods can detect the exact number of SRm160 molecules bound pre-mRNA substrates at single molecule level and (2) they are the only way to reveal any hidden information from bulk measurements. The limitation regarding our findings is that the complexes need to be localised on a surface; thus, there is non-specific binding, which in some cases obscures the binding state and hence a key part of this research has been to validate non-subjective analysis techniques.

In this thesis, the proposed methods are aimed at providing robust results for the state of binding of SRm160 during splicing. The main contributions are summarised in details as follows:

- a. We propose a novel single molecule technique which is introduced as single molecule colocalization system. This system is using a suitable microscope for excitation via total internal reflection (TIRF) to obtain the single molecule data. We also propose this system by using two different fluorophores (the chosen fluorophores are mEGFP and Cy5) that allows us to investigate the interaction of two distinct components in a reaction of interest (GFP-SRm160 and Cy5-labelled-pre-mRNA substrate in a splicing reaction). Multiple conditions are applied, such as the absence and the presence of ATP and other factors. Our experiments consist of various processes. For all these processes, an automatic system

## 1.14 Thesis Outline

---

was available to reduce the time required to achieve one experiment and to increase the consistency across experiments. Additionally, we propose a development procedures regarding data analysis and documentation. Overall, the optimisation via this system ensures a reliable identification of colocalization and bleaching steps.

- b. Motivated by the previous contribution which identified the state of binding of SRm160 during splicing reaction. We propose FRET system to look at the state of SRm160 from another scope. This system allows a deeper investigation of a molecule of interest in a real time, such as, the state of fluctuations of a molecule of interest (forked DNA or double-labelled RNA) in the presence of SRm160. This system is also using a suitable microscope for excitation via total internal reflection (TIRF) to obtain the single molecule data. The chosen fluorophores are Cy5 and Atto647N. An automatic system was also available to reduce the time required to achieve one experiment. Additionally, we propose a development procedures regarding data documentation.

## 1.14 Thesis Outline

The thesis is organized into 8 Chapters. The individual chapters of this thesis are structured as follows:

- Chapter 1 provides an introduction of RNA processing, its biogenesis pathway, the state of SRm160 (the star of the thesis) and an overview of single molecule techniques.
- Chapter 2 provides the material and procedures of popular methods and

## 1.14 Thesis Outline

---

techniques used in this research.

- Chapter 3 presents the fundamental steps of microscope alignments. The performance of data presentation and the analysis procedure are presented. An improvement in the analysis and in the data documentation steps are included.
- Chapter 4 presents specific optimisation of conditions for single molecule experiments.
- Chapter 5 presents observation framework for SRm160 during splicing using single molecule colocalisation technique. Experimental results are presented. The behaviour of SRm160 are analysed. The performance of it is compared to related studies.
- Chapter 6 presents observation framework for SRm160 during splicing using FRET technique. Experimental results are presented. The behaviour of SRm160 are analysed. The performance of it is compared to related observed results via single molecule colocalisation method.
- Chapter 7 concludes the thesis.
- Chapter 8 indicates appendices.

## **Chapter 2**

# **Material and Methods**

# Contents

2.1	<i>In vitro</i> Techniques . . . . .	77
2.1.1	Polymerase Chain Reaction (PCR) . . . . .	77
2.1.1.1	The PCR Procedure . . . . .	77
2.1.1.2	DNA Purification . . . . .	77
2.1.2	<i>In vitro</i> Transcription . . . . .	78
2.1.2.1	The procedure of <i>in vitro</i> Hot transcription . . . . .	78
2.1.2.2	The procedure of <i>in vitro</i> Cold Transcription . . . . .	79
2.1.2.3	The procedure of <i>in vitro</i> with 5GMPS CAP and Labelling with Cy5-maleimide . . . . .	79
2.1.3	<i>In vitro</i> Splicing . . . . .	80
2.1.3.1	The procedure of <i>in vitro</i> splicing . . . . .	80
2.1.3.2	Proteinase PK Treatment . . . . .	81
2.1.3.3	Denaturing Polyacrylamide Gel Electrophore- sis . . . . .	81
2.1.4	Analysis of spliceosomal complexes . . . . .	82
2.1.4.1	Native Agarose Gel Electrophoresis . . . . .	82
2.1.5	Western Blot Analysis . . . . .	82
2.1.5.1	The Procedure of WB . . . . .	82
2.1.5.2	SDS-PAGE Gel Electrophoresis . . . . .	83

## CONTENTS

---

2.2	Oligonucleotide design . . . . .	84
2.3	Cell Culture and Nuclear Extract Preparation . . . . .	84
2.3.1	Transfection of HEK 293T Cell . . . . .	84
2.3.2	Nuclear Extract Preparation . . . . .	85
2.4	Single Molecule Techniques . . . . .	86
2.4.1	Labelling of the pre-mRNA for colocalisation purposes	86
2.4.2	Chamber Preparation . . . . .	86
2.4.3	Preparation of Samples for the splicing complexes . . .	87
2.4.4	Preparation of Samples for the single molecules experi- ments; colocalisation . . . . .	87
2.4.5	Preparation of Samples for the single molecules experi- ments; FRET . . . . .	88
2.4.6	Samples used for FRET experiments . . . . .	88

## 2.1 *In vitro* Techniques

---

## 2.1 *In vitro* Techniques

### 2.1.1 Polymerase Chain Reaction (PCR)

#### 2.1.1.1 The PCR Procedure

All DNA constructs were amplified from plasmids via polymerase chain reaction (PCR) using Red Taq DNA Polymerase (Sigma-Aldrich) following the manufacturer's instructions. This experiment was performed by using G-Storm Thermal Cycler (G-Storm). A 50  $\mu$ l volume reaction contained 50 ng of pure plasmid, 1  $\mu$ l of each forward and reverse primers, 1x buffer, 0.4 mM dNTPs (Promega) and 1 unit of DNA polymerase (NEB, UK). Cycle conditions were set up according to manufacturers recommendations. Annealing temperature varied depending on the oligonucleotide sequence composition. PCR was done during 35 cycles. The PCR reaction was analysed by running on a 1% Agarose gel (Helena) to ensure the amplified fragment was at the correct size and to confirm the presence of PCR products. Loading buffer, 6x (NEB) was used in appropriate proportions to the PCR reaction in 1x TAE buffer. 2 l of 1 kb DNA ladder (NEB) were used to determine the size of DNA fragments Ethidium bromide (EtBr, 0.01%) was involved in the Agarose Gel for impurities. Electrophoresis was performed at 80 V for about 45 min. PCR products were analysed under UV light. In the end, products were excised extracted with phenol-chloroform and ethanol precipitated. Their concentrations were determined by measuring the absorbance (A260) using nanodrop.

#### 2.1.1.2 DNA Purification

Extraction of DNA via phenol-chloroform was done after the PCR reaction; autoclaved ultra-pure water was first added up to a final volume of 100  $\mu$ l.

## 2.1 *In vitro* Techniques

---

The same amount of Phenol: chloroform: Isoamyl Alcohol 25:24:1 (Sigma-Aldrich) was added to the samples, mixed by vortexing and centrifuged for 3 min at 13000 rpm. The upper aqueous phase was collected in a fresh tube and this step was repeated. 100  $\mu$ l of chloroform (Sigma-Aldrich) was added to the sample in a fume hood, mixed by vortexing and spun under the same conditions. The supernatant was removed. 80% ethanol was added followed by centrifugation for 5 min. The supernatant was removed and the volume of the sample was assessed. 1/10 volume of 3M NaAc pH 5.2 and 2 volumes of 100% ethanol were added and mixed. Samples were centrifuged at 17,000 rpm for 15 minutes. The supernatant was removed. The sample was dried under vacuum for 20 min. The DNA pellet then re-suspended in 20  $\mu$ l TE.1 buffer. The samples were stored at  $-20^{\circ}\text{C}$ .

### 2.1.2 *In vitro* Transcription

#### 2.1.2.1 The procedure of *in vitro* Hot transcription

This procedure is introduced as 'Hot transcription'. DNA fragments containing the T7 pro-moter at the 5' end were generated by PCR (Section 1.1.1). The homemade T7 RNA polymerase was used to generate transcripts for *in vitro* splicing assays. Typically, a 10  $\mu$ l of transcription reaction containing transcription buffer (40 mM Tris-HCl pH 7.5, 6 mM MgCl<sub>2</sub>, 10 mM NaCl), 0.5 mM ATP, 0.5 mM CTP, 0.5 mM UTP, 0.05 mM GTP (Promega), 50 ng/  $\mu$ l DNA template, 10 U RNaseOut (Invitrogen), 5% T7 RNA polymerase (1:20), 1 mM Ribo m7G cap analogue (Promega) and 0.5  $\mu$ M [<sup>-32</sup>P] GTP (10 mCi/ml, 3000 Ci/mmol; Pelkin Elmer) was set up on ice and then incubated for 2 h at 37°C. After incubation, samples were mixed with 10  $\mu$ l formamide dyes (90 % v/v formamide, 50 mM EDTA, bromophenol blue and xylene cyanol)

## 2.1 *In vitro* Techniques

---

and fractionated on 6% denaturing polyacrylamide gel electrophoresis. The transcripts was visualized using x-ray film (Fujifilm) and excised with a clean scalpel. The transcripts was eluted into elution buffer (1 mM EDTA pH 8, 0.2% SDS and 0.5 M NaAC pH 4) and stored overnight at 4°C. In the following day, the eluted solution was collected into a clean eppendorf, ethanol precipitated, re-suspended in TE.1 buffer and finally incubated at -80°C.

### 2.1.2.2 The procedure of *in vitro* Cold Transcription

Transcription reaction of final volume of 50  $\mu$ L containing 40 mM Tris-HCl pH 7.5, 6 mM MgCl<sub>2</sub>, 10 mM NaCl, 0.5 mM of each rNTPs (0.5 mM ATP, 0.5 mM rCTP, 0.5 mM rUTP, 0.05 mM rGTP), 15 ng/ $\mu$ L linear PCR-amplified DNA template (or 160 ng/ $\mu$ L of digested plasmid), 10 U/ $\mu$ L RNase OUT (Invitrogen), and 5% T7 RNA polymerase (1:20) and 1 mM Ribo m7G cap analogue (Promega) was set up on ice and incubated for 4 h at 37°C. One unit DNase (Promega) was added and incubated for 30 min at 37°C to degrade remaining DNA. One unit per  $\mu$ g of DNase (Promega) was added to the reaction and it was incubated for 10 min at 65°C to stop the enzymatic activity. The reaction was subsequently made up to 100  $\mu$ L with water. The transcript was then extracted via phenol-chloroform. An S-300 column (GE Healthcare) was used to purify the RNA. The RNA was phenol-chloroform extracted and ethanol precipitated before being re-suspended in 30-50  $\mu$ L TE.1 buffer in addition of 10 U RNaseOut. RNA was stored at -80°C.

### 2.1.2.3 The procedure of *in vitro* with 5GMPS CAP and Labelling with Cy5-maleimide

A 50  $\mu$ L cold transcription reaction containing 40 mM Tris-HCl pH 7.5, 20 mM MgCl<sub>2</sub>, 10 mM NaCl, 10 mM DTT, 4 mM ATP, 4 mM CTP, 4 mM UTP, 1

## 2.1 *In vitro* Techniques

---

mM GTP, 15 ng/ $\mu$ l linear PCR amplified DNA template (160 ng/ $\mu$ l if digested plasmid is being added), 10 mM Guanosine- 5'- O- mono phosphorothioate (5-GMPS) (BioLog G 018), 2 U/ $\mu$ l RNase OUT (Invitrogen) and 6% homemade T7 RNA polymerase (1:20) was set up on ice and incubated for 4 h at 37°C. 1 unit per  $\mu$ g of DNase (Promega) was added to the reaction and incubated for 15 min at 37°C. The reaction was subsequently made up to 100  $\mu$ l with water and extracted with phenol-chloroform. An S-300 column (GE Healthcare) was used to purify the RNA. The RNA was then ethanol precipitated and re-suspended in 17  $\mu$ l H<sub>2</sub>O, 2  $\mu$ l NaPO<sub>4</sub> pH 7.2 and 1  $\mu$ l of 10 mM Cy5-maleimide in DMSO (Cy5-maleimide) (Lumiprobe 13080). The reaction was incubated at room temperature for 4 hours. The Cy5-maleimide labelled RNA was then gel purified. The bands were visualised using UV shadowing (see the section below).

### 2.1.3 *In vitro* Splicing

#### 2.1.3.1 The procedure of *in vitro* splicing

*In vitro* splicing is typically carried out in 10  $\mu$ l reaction containing a suitable amount of radiolabelled RNA, 1.5 mM rATP, 20 nM HEPES pH 7.5, 20 mM of C-reactive protein inhibitors (CrPi), 3.2 mM MgCl<sub>2</sub>, 50 mM potassium glutamate (KGLu) and 50% nuclear extract. The reaction was set up on the ice and incubated at 30 °C for 2 h. At designated timepoints (at 0, 30, 60, 90, and 120 min time points), aliquots of 2  $\mu$ l of RNA were taken into a microtitre plate on dry ice. Each aliquot was subsequently treated with Proteinase K (PK), see the next section. Samples were then ethanol precipitated and dissolved in 10  $\mu$ l of F-dye (containing 10 mM EDTA pH 8 and 0.2% Xylene Cyanol and 0.2% Bromothol Blue). Samples were heated at 80°C for 30 seconds then

## **2.1 *In vitro* Techniques**

---

run on a 6% denaturing polyacrylamide gel. The gel was subsequently dried and exposed to a phosphorimaging screen. Quantification of both pre-mRNA and mRNA bands was done using OptiQuant software (PerkinElmer).

### **2.1.3.2 Proteinase PK Treatment**

This treatment was done using a mixture of 50  $\mu$ l Proteinase K buffer (100 mM Tris-HCL pH 7.5, 12.5 mM 0.5 M EDTA, 150 mM NaCl, 1 % SDS) and 2  $\mu$ l of 10 mg/ ml proteinase K in 10 mM CaCl<sub>2</sub>. This treatment was done at 37°C for 15 min. 100  $\mu$ l of 80 % ethanol was added followed with centrifugation at 62000 rpm for 15 min (Rotonta 460 R). The liquid was discarded using water pump. RNA was washed with 150  $\mu$ l 100 % ethanol and spun again. 10  $\mu$ l of formamide dyes were added to each sample prior loading on denaturing polyacrylamide gel.

### **2.1.3.3 Denaturing Polyacrylamide Gel Electrophoresis**

20 ml and 40 ml gels contained 7 M Urea, 6 % acrylamide 19:1 bisacrylamide solution (National Diagnostics) and 1X TBE buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA). Polymerisation was achieved using 0.1% ammonium persulfate (AMPS) and 0.002% TEMED (Sigma). 10  $\mu$ l of Formamide dyes were added to samples prior loading onto the gel. The electrophoresis was carried out for various times depending on the length of RNA. Splicing gels were dried and exposed to Phosphor screens (Packard) for analysis.

## **2.1 *In vitro* Techniques**

---

### **2.1.4 Analysis of spliceosomal complexes**

#### **2.1.4.1 Native Agarose Gel Electrophoresis**

Splicing reactions were set up on ice and 2  $\mu\text{l}$  of each time course was taken and frozen on dry ice, see section 2.2.5. For A, B, and C complex, a 1.75  $\mu\text{l}$  mix containing 0.5  $\mu\text{l}$  of 4 mg/ml heparin and 1.25  $\mu\text{l}$  of TG loading dye (50 mM Tris, 50 mM glycine 40% glycerol, bromophenol blue, and xylene cyanol) was added. Samples were then run on a 2% native UltraPure Low Melting Point Agarose (Invitrogen) gel (50 mM Tris, 50 mM glycine) at 100 V for 4 h to 5 h at 4°C. For H and E complex, no heparin was used and only 1  $\mu\text{l}$  of TG loading dye added before running on a 1.5% gel. The gel was then compressed between two paper sheets (3 mm chromatography paper (Whatman)) surrounded by multiple layers of paper towel to remove excess liquid and make the gel thin enough. Afterwards, the gel was dried and exposed to a phosphorimaging screen.

### **2.1.5 Western Blot Analysis**

#### **2.1.5.1 The Procedure of WB**

After protein samples have been transferred to nitrocellulose, the membrane was blocked in blocking buffer (5% Milk/ BSA, 0.1% Tween-20, 1xTBS (150 mM NaCl, 10 mM Tris pH8.0)) for 1 h at room temperature. It was then incubated for another 1 hour at 4°C in a relevant dilution of primary antibody in blocking buffer. The membrane was then washed 5 times (5-10 min per time) in TBST buffer (0.1% Tween-20, 1xTBS) before a secondary antibody was applied (in blocking buffer at 1:1000 dilution) to the membrane for 1 h at 4°C. Washing steps were repeated again after the secondary antibody.

## 2.1 *In vitro* Techniques

---

The membrane was either developing with ECL reagents and then exposed to X-ray film (Fuji) or scanned using a fluorescent scanner.

### 2.1.5.2 SDS-PAGE Gel Electrophoresis

All SDS-PAGE gels were made of 80% volume of resolving gel at the bottom (390 mM Tris pH 8.8, 8-12% acrylamide (ProtoGel, National Diagnostics), 0.1% SDS polymerized with 1% AMPS and 0.2% TEMED) and about 20% volume of stacking gel (125 mM Tris Base pH 6.8, 4% acrylamide, 0.1% SDS polymerized with 0.1% AMPS and 0.2% TEMED) at the top with an appropriate size comb. Protein samples were heated at 100°C for 10 min in SDS-PAGE loading dye (50 mM Tris pH 6.8, 2% SDS, 10% Glycerol, 100 mM DTT, and 0.05% Bromophenol blue) before loading. The gel was run in 1X SDS-PAGE running buffer (125 mM Tris Base, 192 mM Glycine, 0.1% SDS) at 80 V until the samples were in the resolving gel part and at 120 V for the remainder. A protein marker (Bio-Rad) and Bromophenol blue, in the loading dye were used to track the progression of the samples. Gels were then stained or transferred to nitrocellulose as required. SDS-PAGE gels were soaked briefly in 1X transfer buffer (25 mM Tris, 192 mM Glycine, 10% Methanol) and placed onto a nitrocellulose membrane (BioTrace). The nitrocellulose and gel are placed between 2 stacks of 5 pieces of 3 mm chromatography paper (Whatman) (damp with transfer buffer). Transfer buffer was added at each stage of assembly to prevent bubbles in the stack. The transfer was done by running (Biometra Fast-Blot B33) at 10 W for 30 min.

## 2.2 Oligonucleotide design

---

## 2.2 Oligonucleotide design

Oligonucleotides were designed to be about 50 bp in length. All oligonucleotides used in this section are designed from GloC. The insert complementary part was designed to ensure an annealing temperature to be around 55-60°C. Oligonucleotides were amplified via PCR and then tested via Agarose gel electrophoresis to ensure the right size of each oligonucleotide as we planned. They were dissolved in TE.1 buffer (10 mM Tris pH 7.5, 0.1 mM EDTA pH 8) to stock concentrations of 100 M and stored at -20°C. All primers and oligonucleotides used are listed in Chapter 8, section 8.1.12 to 8.1.22.

## 2.3 Cell Culture and Nuclear Extract Preparation

### 2.3.1 Transfection of HEK 293T Cell

Cells were diluted in DMEM medium (Invitrogen) supplemented with 10% FBS (Gibco) and 1% Penstrep and seeded onto 15 cm plates (Corning) with a density of  $5 \times 10^4$  cells per plate. Cells were allowed to grow overnight then the medium was changed 4 h before beginning calcium chloride transfection. The desired amount of plasmid (8-60  $\mu$ g) in TE.1 buffer was mixed on ice with 10x volume of CaCl<sub>2</sub> solution (1 mM Tris-HCl pH 7.5, 0.1 mM EDTA, 300 mM CaCl<sub>2</sub>). An equal volume of pre-chilled HBS solution (342.23 mM NaCl, 12.41 mM KCl, 1.76 mM Na<sub>2</sub>HPO<sub>4</sub>, 13.88 mM glucose, 52.45 mM HEPES pH 7.5) was added and incubated on ice for another 10 min. The mixture (about 3.5 ml) was carefully added to the plate drop-wise and incubated for 16-24 h. Cells were then shocked for 3 min with 20 ml of 25% DMSO in DMEM

## 2.3 Cell Culture and Nuclear Extract Preparation

---

solution and washed twice with 15 ml DMEM. This was then incubated for 16-48 h in 20 ml fresh complete medium before harvest.

### 2.3.2 Nuclear Extract Preparation

After transfection step, cells were washed once with 15 ml cold PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4). Cells were harvested with a cell scraper in 10 ml cold PBS. The cells were centrifuged at 150 g for 8 min. The cell pellet was then washed with 1 ml of PBS and transferred into a 1.5 ml tube and pelleted by centrifugation at 4°C at 9250 g for 5 min. The packed cell volume (PCV) was washed and re-suspended in one PCV of buffer A (10 mM HEPES pH 8, 1.5 mM MgCl<sub>2</sub>, 10 mM KCl, 1 mM DTT, 0.06% NP40) and incubated for 15 min on ice to swell the cells. After swelling, cells were lysed by vortexing for 15 seconds. The lysed cells were centrifuged at 9250 g at 4°C for 1 min. Nuclei were pelleted by centrifugation at 13,000 rpm for 30 seconds at 4°C. The supernatant was removed and 0.7 PCV of buffer C (20 mM HEPES pH 8, 25% glycerol, 420 mM NaCl, 0.2 mM EDTA, 1 mM DTT) added. The sample was incubated for 30 min on ice with mixing on a magnetic stirrer using a small magnetic flea. Debris was removed by centrifugation at 13,000 rpm for 10 min at 4°C. Nuclear extract supernatant was dialyzed against buffer D (10% glycerol, 20 mM HEPES pH 8, 0.2 mM EDTA, 1 mM DTT, 100 mM KCl), on a 0.025 μm VSWP membrane filter (Millipore) at 4°C for 2 h and snap-frozen with liquid nitrogen in aliquots and stored at -80 °C. The dialysed nuclear extract was centrifuged for 5 min at 9250 g and stored at 80°C. All nuclear extracts were tested to ensure its activity. *in vitro* splicing experiments were the platform to do this test in order to show sufficient splicing efficiency

## 2.4 Single Molecule Techniques

---

### 2.4 Single Molecule Techniques

#### 2.4.1 Labelling of the pre-mRNA for colocalisation purposes

Pre-mRNA transcripts were labelled by annealing an oligonucleotide containing a fluorophore Cy5 (See Section 8.1.23) to them. Reactions were set up in 100 mM NaCl, 10 mM HEPES pH 8, 1.67  $\mu$ M pre-mRNA and 1.67  $\mu$ M of the Cy5 oligo. PCR machine was used to implement the annealing process. 1  $\mu$ l of each sample were run with 4  $\mu$ l H<sub>2</sub>O and 5  $\mu$ l Formamide containing 10 mM EDTA pH 8 on a native 6 % acrylamide gel and imaged by the Typhoon (GE Healthcare).

#### 2.4.2 Chamber Preparation

Glass cover slides, 22 x 50 mm #1 (Menzel-Glser) were incubated in 1M KCL, washed with distilled water, sonicated in a water bath for 15 min and dried under an N<sub>2</sub> stream. They were then cleaned in an argon plasma (MiniFlecto-PC-MFC, Gala Instrument) 5 times for 5 min with pure Argon at 0.15 mbar and an applied power of 80 W. Double-sided tape Duplocoll 370 (Lohmann Adhesive Tapes) was used to create a 5 mm to 10 mm wide channel parallel to the cover slide. The channel was covered with another glass cover slip of size 22 x 22 mm #1.5 in order to form the sample chamber. By this stage, the slide chamber is applicable to use for single molecule colocalisation experiments. In the case of FRET experiments, the sample chamber was modified with Biotin-BSA and streptavidin (section 1.3.5).

## 2.4 Single Molecule Techniques

---

### 2.4.3 Preparation of Samples for the splicing complexes

Splicing reactions stalled at different complexes were prepared for single molecule microscopy. All samples were prepared with 50 % nuclear extract, 3.2 mM MgCl<sub>2</sub>, 50 mM KGlu and 1 units RNase OUT (Invitrogen). This was used for complex E, as it did not contain any ATP. progression splicing to past complex E was achieved by the addition of 1.5 mM rATP, 20 mM CrPi, and 20 mM HEPES pH 7.5. anti-U6 oligo at 1  $\mu$ M and incubation at 30°C for 5 min were additionally performed to halt splicing at complex A. To block the U1 snRNP, 3.3 $\mu$ M of U1 oligo was added. Similarly, U2 oligo was added to the splicing reaction to block U2 snRNP. For complex C, the Globin C GG pre-mRNA was added where the 3 splice site sequence was replaced by a random sequence. The splicing reaction was made without RNA and pre-incubated for 15 min at 30°C. Then, Cy5-labelled RNA was added to the splicing reaction for a final concentration of 62.5 nM and incubated for further 15 min at 30°C.

### 2.4.4 Preparation of Samples for the single molecules experiments; colocalisation

Once the complex is formed, serial dilutions of the samples were prepared down to a concentration of 5 pM of the pre-mRNA using splicing buffer (buffer A) which contains (3.2 mM MgCl<sub>2</sub>, 50 mM Potassium Glutamate, 50% Buffer D, and 0.04 U/ $\mu$ l RNase OUT). This was to optimize the most suitable density of spots for microscopy. The highest diluted samples were loaded into the sample chamber and incubated for 5 min. The density of the apparent spots was judged, and, if considered too low, a higher concentrated sample was loaded.

## 2.4 Single Molecule Techniques

---

### 2.4.5 Preparation of Samples for the single molecules experiments; FRET

For FRET purposes, the sample chamber was washed with T50 buffer (ingredients) for purities, incubated with 20  $\mu\text{g}$  biotin-BSA in PBS for 10 min, washed with T50 buffer buffer A (100 mM NaCl, 50 mM HEPES pH 7.5, 1  $\mu\text{M}$  DTT, 20 units/ml RNase OUT (Invitrogen)), incubated with 10  $\mu\text{g}$  streptavidin (Invitrogen) in PBS and again washed with T50 buffer (buffer A). The sample of interest was applied at various concentrations depending on the experiment of interest. A final wash with T50 buffer to ensure there is no excess of the free sample remain. Scavenger and anti-blinking reagents were used when needed.

### 2.4.6 Samples used for FRET experiments

Forked DNA was used as control experiments to optimize the system and to possess suitable conditions. This construct was a double labelled DNA with two fluorophores; the acceptor fluorophore (Cy5) and the donor fluorophore (Cy3). Those fluorophores were incorporated within the DNA. The biotinylated site was introduced to allow the substrate to immobilize to the surface (sample chamber), see (Chapter 6, section 6.2). Additionally, a double-labelled RNA was the sample of interest. It was a kind gift from Prof. Eperon. This substrate contains two fluorophores; the acceptor fluorophore (ATTO 247N) and the donor fluorophore (Cy3). Fluorophores were incorporated within the DNA. A biotin site was attached to the 3'-end allowing the attachment of the RNA to the slide chamber. FRET experiments were performed in the absence and the presence of an active commercial nuclear extract. Moreover, scavenger reagents were implemented to prevent photobleaching.

## Chapter 3

Using of The Microscope: *Set up, alignment, data presentation and analysis*

# Contents

3.1	The Microscope Set up . . . . .	91
3.2	The Two Lasers Fluorescence Microscopy . . . . .	95
3.3	The Camera Vs. the Detection Efficiency . . . . .	96
3.4	The Sample Stability . . . . .	97
3.5	Data Acquisition via LabView . . . . .	98
3.6	Composite Images . . . . .	100
3.7	Spot Detection . . . . .	104
3.8	Inspection of Colocalization . . . . .	104
3.9	The Assignment of Steps . . . . .	108
3.10	Analyzing Data . . . . .	108
3.11	Presentation of Data . . . . .	110

### 3.1 The Microscope Set up

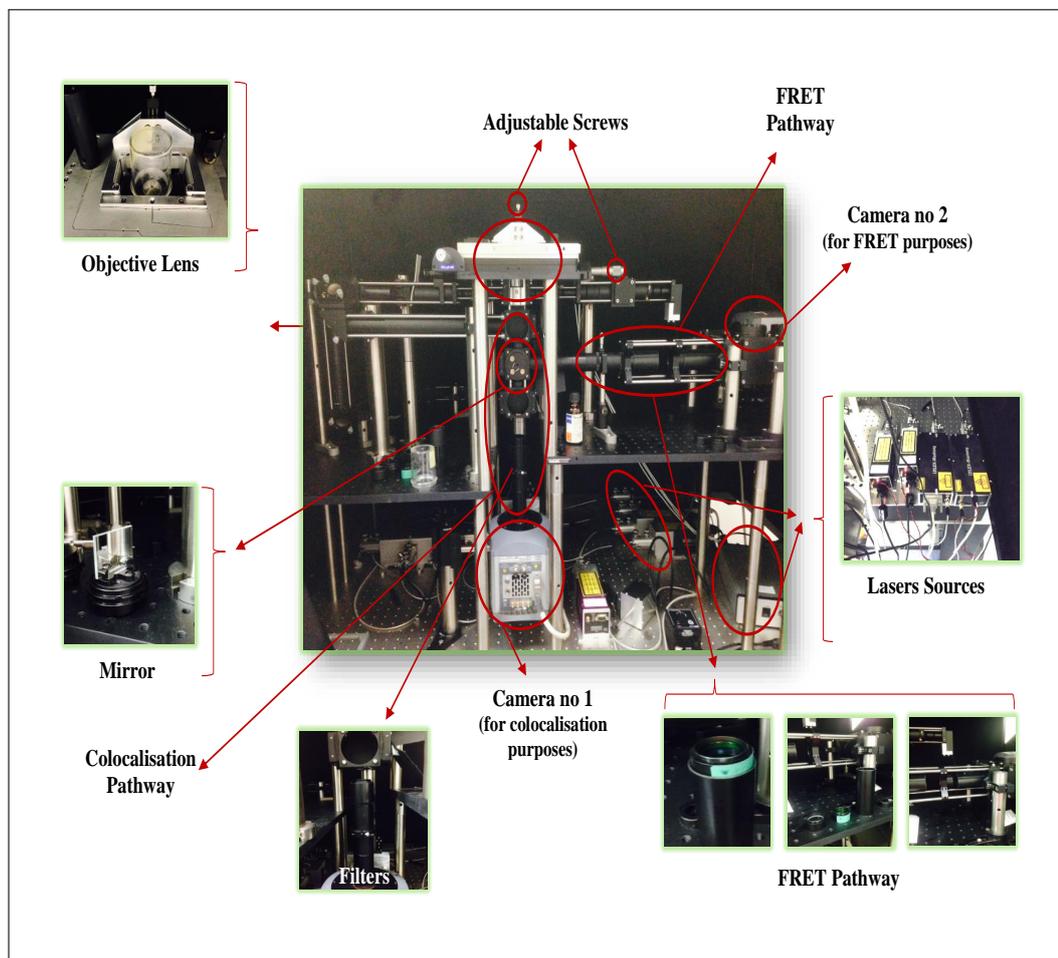


Figure 3.1: A complete view of our home made single molecule microscope. The most important parts are labelled. The details of these parts are represented later in Figure 3.2 and 3.3

### 3.1 The Microscope Set up

A homemade fluorescence microscope was used in this work. The instrument was built by Dr. Andrew Hudson (Figure 3.1). Before doing any experiment and after the single molecule sample is prepared, the microscope needs to be optimized to ensure the possible accurate results.

For our research, the microscope was adjusted to be suitable for excitation via total internal reflection (TIRF) to obtain the single molecule data.

### 3.1 The Microscope Set up

---

Figure 3.2 and 3.3 shows the main parts of the microscope. Each part is labelled in red.

Multiple lasers are available for the wavelengths of 488 nm, 532 nm, 561 nm and 633 nm. These beams were generated from separate laser diode or diode-pumped solid-state lasers (1). They fed into a the beam combiner (2) that aligns all laser beams along the same beam path. The combiner was connected to the microscope via an optical fibre. The merged lasers were plane polarised and passed through different objects, from rotatable wave-plates, splitters, multiple mirrors and finally steered onto the objective lens (Figure 3.2 and 3.3).

The plane of polarisation of the laser could be rotated via a half wave-plate (3). The polarising beam splitter (4) separates the incoming beam into two beams with orthogonal polarisation, of which the transmitted horizontally polarised beam was used. A quarter wave-plate changes the polarisation of the laser beam to a circular polarised (5). Then, the laser beam was widened by passing through a beam expander (Keplarian telescope) which changes the beam diameter from 1 mm to 20 mm, (7) and (8). The diaphragm (10) blocks light that does not pass through its aperture of a diameter greater than 7 mm. As the laser beam has a Gaussian intensity profile before the aperture, widening the beam width provides a more homogeneous intensity profile.

The lens (11) focusses the laser beam onto the back focal plane of the objective. This lens leads to a collimated laser from of the objective lens at the interface to the sample (13). The lasers were directed using multiple adjustable mirrors which allowed the beam to be aligned onto the edge of the back aperture of the objective (6) and (12). The laser beam then reaches the sample interface at an angle greater than the critical angle for total internal

### 3.1 The Microscope Set up

---

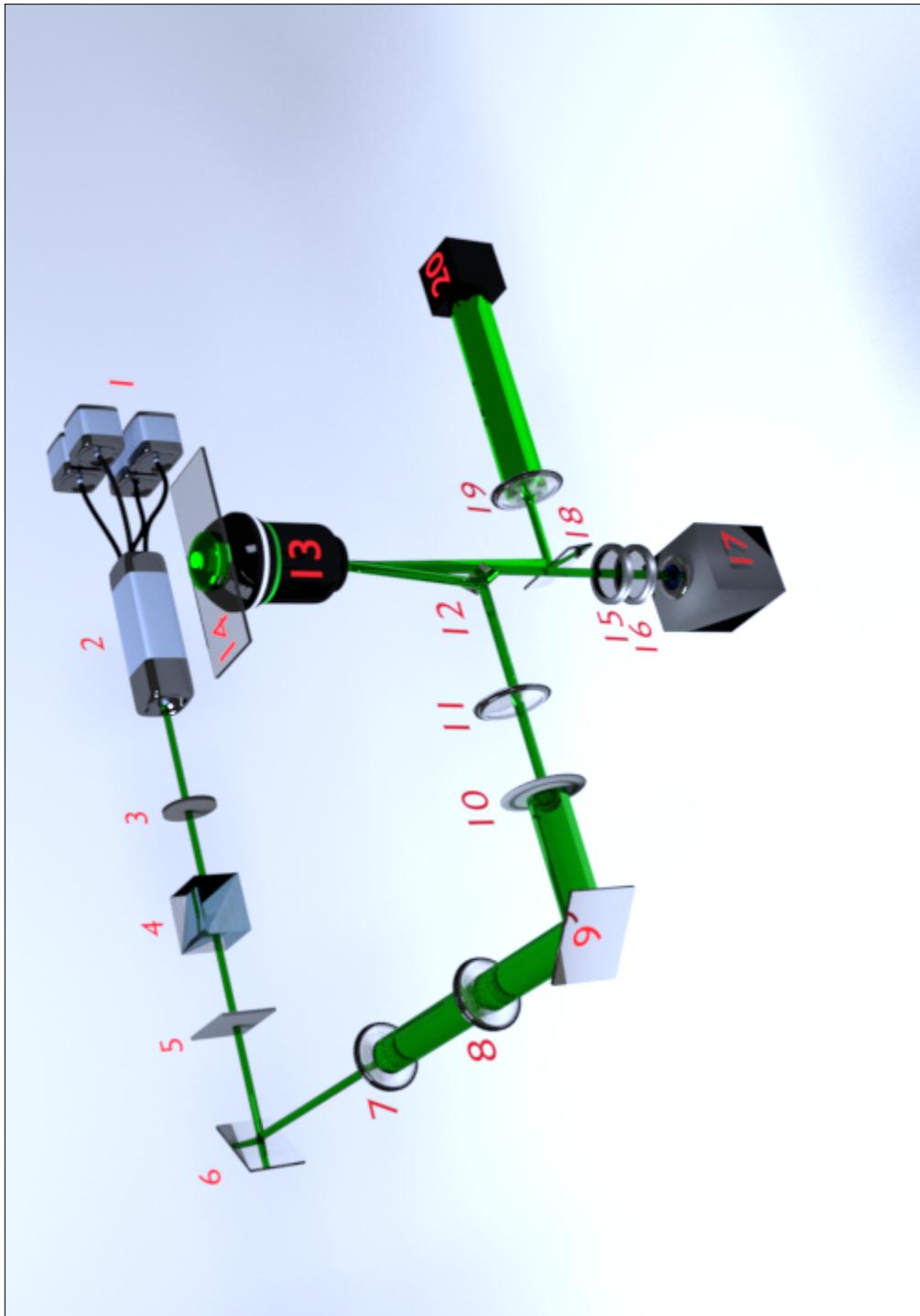


Figure 3.2: Schematic diagram of our home made single molecule microscope. Each part is numbered in red. Those parts indicate the main set-up used for single molecule colocalisation measurements.

### 3.1 The Microscope Set up

---

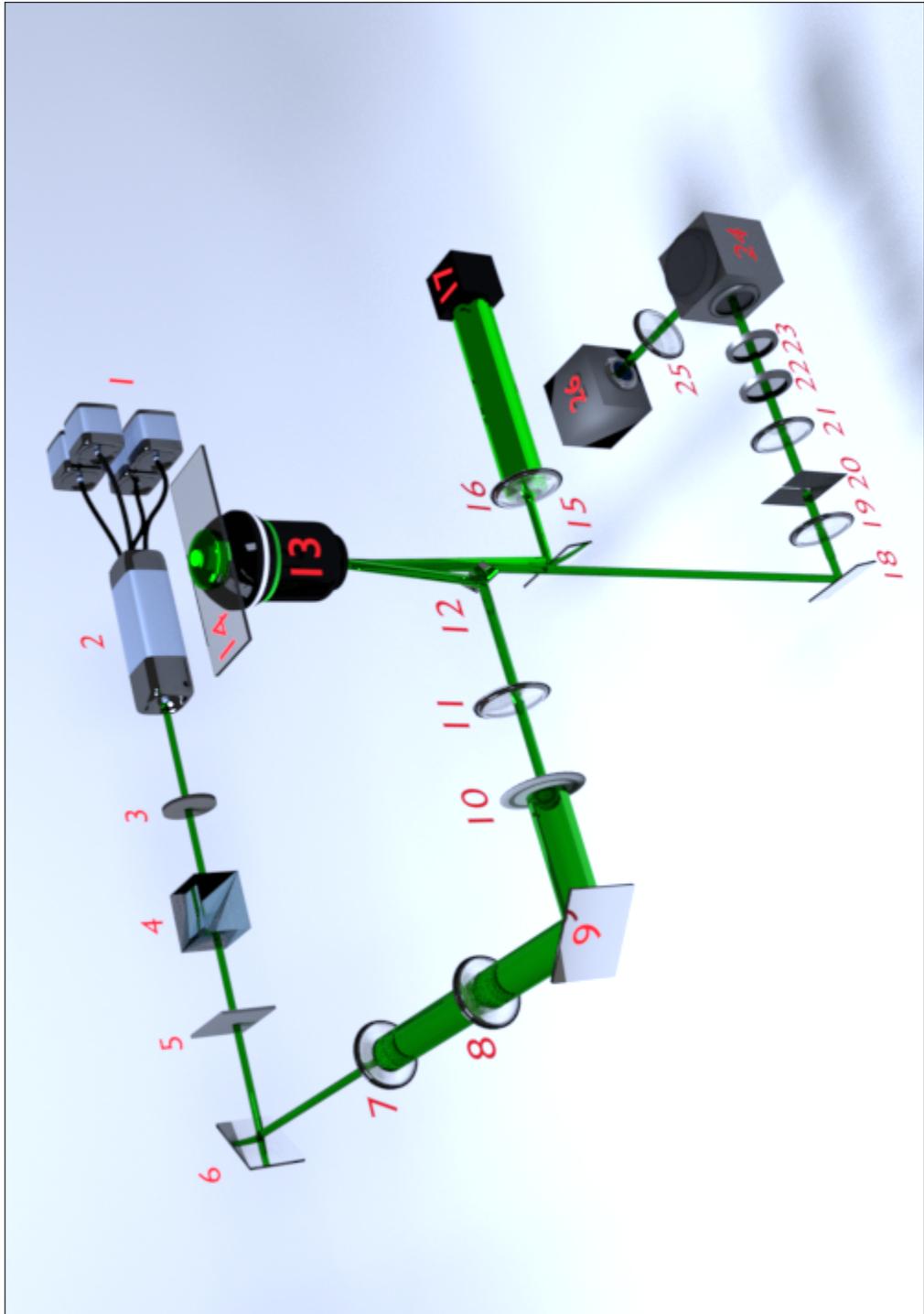


Figure 3.3: Schematic diagram of our home made single molecule microscope. Each part is numbered in red. Those parts indicate the main set-up used for single molecule FRET measurements.

### 3.2 The Two Lasers Fluorescence Microscopy

---

reflection. At this stage, total internal reflection occurs, and an evanescent wave penetrates the sample and excites it (14). The fluorescent emission is collected by the same objective and passes through notch filters suited for the wavelengths of the laser beams used for the excitation (15) and (16) and a tube lens. EMCCD to records the image (17).

In the situation of using the total internal reflection fluorescence microscopy (TIRFM) for Fluorescence resonance energy transfer (FRET) purposes, parts from (1) to (14) as same as parts explained previously in the colocalisation set up. Parts (15) to (17) are corresponding to the parts from (18) to (20) in the colocalisation set up. Those are responsible for the laser auto-focusing step. The beam is directed onto a quadrant photodiode (17) and used to keep the sample in focus. The differences between the two paths are: The presence of rotatable mirror (18). This mirror is easily adjusted to reach the right path (between the two cameras; (17) for the colocalization path (Figure 3.2) and ((26) for the FRET Path (Figure 3.3)). The florescence is passed through a lens (19) to concentrate the beam, adjustable slit (20) to another lens (21). Two specific filters (22) and (23) were used depending on the laser's wavelength used in the experiment. A correction filter is also used to prevent any other wavelengths to be collected. Parallel dichroic mirrors (24) were used. Each mirror is a specific for a specific fluorophore. Finally the florescence light is passes through a lens (25) to EMCCD.

### 3.2 The Two Lasers Fluorescence Microscopy

The use of two different fluorophores allows us to investigate the interaction of two distinct components in a reaction of interest. The chosen fluorophores

### 3.3 The Camera Vs. the Detection Efficiency

---

in this research are mEGFP and Cy5. These fluorophores are suitable for the colocalisation system. On the other hand, in the FRET system, the chosen fluorophores are Cy5 and Atto647N. They all provide a sufficient spectral separation.

As mentioned previously in section 3.1, our home made system provides lasers beam of the wavelength 488 nm, 532 nm, 561 nm and 633 nm. Of those, the wavelengths of 488 nm, 532 nm and 633nm are used in this research. They used to excite mEGFP, Cy3, and Atto67N or Cy5 respectively.

For the single molecule colocalization experiments, it was decided to use sequential illumination. The longest wavelength was used for the excitation until all corresponding fluorophores have bleached. Thereafter, the shorter wavelength has applied. This procedure allows us to obtain images the different fluorophores.

In our setup, an excitation beam with the 633 nm has used first to excite Cy5. This wavelength is unable to excite GFP such that the collected fluorescence signal has obtained only from Cy5 fluorophores. This was done until all Cy5 fluorophores are bleached. Then the excitation was switched to the wavelength of 488 nm for the final excitation of GFP. By this stage, the collected fluorescence signals originate only from GFP fluorophores. On the other hand, in the case of FRET experiments, an excitation beam with the 633 nm was used to excite donor fluorophores (Atto647N).

### 3.3 The Camera Vs. the Detection Efficiency

However, optimising the setup of the microscope ensures reliable identification of spots and bleaching steps. Various events can affect the legibility of images

### 3.4 The Sample Stability

---

obtained by the microscope. Detection efficiency, lowest background, and noise levels are examples of these events. Decisive for this is the camera and its settings. Our system uses an iXon EMCCD camera from Andor. An important setting for the camera is the exposure time. The longer the exposure time, the more photons can be collected per readout that creates clearer images. Also, this helps the resolution of the acquisition to be good enough. However, bleaching steps are observed at longer times. Exposure times of 0.1 s have been found to provide a good compromise between these two requirements.

The central 250 pixels by 250 pixels part of the 512 pixels by 512 pixel CCD chip has used for all acquisitions recorded during the colocalization experiments. As the illumination with laser beam leads to an unavoidable background which would be homogeneous across the whole field of view. Only the central part of the laser beam has used. This step reduces the background level. The resulting field of view is smaller but with homogeneous intensity distributions.

### 3.4 The Sample Stability

Data acquisitions are done over time. Due to this, the position of the sample should not change during this time. Changes of the position of the sample would move the sample out of the focus of the objective and make the data analysis complicated.

The sample chamber lies on top of a film of immersion oil on top of the objective lens. The immersion oil shows slow relaxation times due to its high viscosity that leads to vertical shifts of the sample chamber out of the focus over time. Given enough time, the system would find an equilibrium with a

### 3.5 Data Acquisition via LabView

---

stable position. Frequent movements of the sample chamber in order to record different areas of it make this option unfeasible. Additionally, the focus shifts for various excitation wavelengths. The image acquisition for each fluorophore and its corresponding wavelength of excitation should be performed in focus. This was achieved using LabView automatically correct the sample position. This was developed by Dr. Robert Weinmeister [Weinmeister, 2014].

### 3.5 Data Acquisition via LabView

One experiment consists of various sequential acquisitions. For each acquisition, the sample has to be moved to image a new region. Generally, the first laser is turned on for a particular time before the illumination is switched to a second laser. The focus has to adjust for each laser during acquisition. The acquired data (the focus) has to be saved. The sample is then moved to a new position in order to start a new record. The process has repeated multiple times until enough data is obtained. For all these processes, an automatic system was available to reduce the time required to achieve one experiment and to increase the consistency across experiments. A LabView program was written by Dr. Robert Weinmeister [Weinmeister, 2014] to control the microscope and to perform the automatic support and recording.

Figure 3.4 shows the LabView interface (an integrated development environment designed for building measurement and control systems). Each part is labelled in green. The encircled numbers correspond to the numbers in the figure. In our system, there are two cameras attached to the microscope, the bottom camera, and the right camera. For the colocalization experiments,

### 3.5 Data Acquisition via LabView

---

the bottom camera was selected. The camera has to be cooled down to  $-80^{\circ}\text{C}$  before it is used (Figure 3.4, (1)). The cooling process was automatically initiated once the program is started. Once the target temperature is reached and stabilized, the rectangle (Figure 3.4, (1) and (2)) becomes green. The exposure time and EM gain of the camera can be chosen (Figure 3.4, (3)). The EMCCD is turned on and set to a 300 unit gain. Also, the camera settings can be set manually as in (Figure 3.4, (6)). The camera chip has 512 pixels by 512 pixels. Only parts of it or multiple individual pixels can be used, such that the acquisition area had to be binned (Figure 3.4, (5)). Each laser can be turned on and off individually, either manually or automatically (Figure 3.4, (4)). Focus points can be saved (Figure 3.4, (4)). If an acquisition is in progress, each recorded image and single molecule spots are shown in (Figure 3.4, (7)). Additional information about the current acquisition is shown in (Figure 3.4, (8)). A quick vertical adjustments for the stage holding the sample can be made with a slide (Figure 3.4 (9)) or can be moved in all three dimensions with sub-micrometre precision (Figure 3.4, (10)). Once the sample position was adjusted and the focus was saved, the output (a small white circle could be seen in the centre) was as shown in (Figure 3.4, (13)). This output was used for the feedback that adjusts the focus automatically once initialized (Figure 3.4 (11)). Each acquisition could be performed manually or automatically. The change of lasers and the end of an acquisition could be set to occur at certain frames. The stage was allowed to move within its limits to repeat the acquisition process nine times, automatically saving the data for each one (Figure 3.4, (14)). The only actions still needed to be performed manually were the initial mounting of the sample and moving the sample to allow the imaging of a new area for every nine acquisitions. Acquisitions were saved as

### 3.6 Composite Images

---

TIFF files. Each file name was automatically complemented with the current date and time to allow a later independent assignment to experiments and to prevent the accidental overwriting of data. Each acquisition was accompanied by a text file that describes the details of each experiment. These details could also be optionally written in the experimental notes (Figure 3.4, (13)).

## 3.6 Composite Images

The acquired data consisted of sequential images (Figure3.5). Information about the wavelengths used for the excitation for each image could be obtained by inspecting the raw data, knowing the experimental procedure and/ or from the accompanying text file. The information from all the images corresponding to one wavelength was used to create composite images in order to reduce the background noise seen in a single image and highlight individual spots.

Composite images were created by computing the mean or the maximum value for each pixel across all images (Figure3.6). Once the maximum values were computed, a spot with a high intensity could more easily be recognised. Computing the mean values helped to recognise a spot with lower intensity emitting fluorescence over the whole acquisition time as its intensity was consistently higher than the surrounding background. In most cases, spots could be identified, but some might have been missed. To avoid this, both maximum and mean values were used in conjunction to reliably detect spots that were either short-lived with a higher intensity or long-lived with a lower intensity.

### 3.6 Composite Images

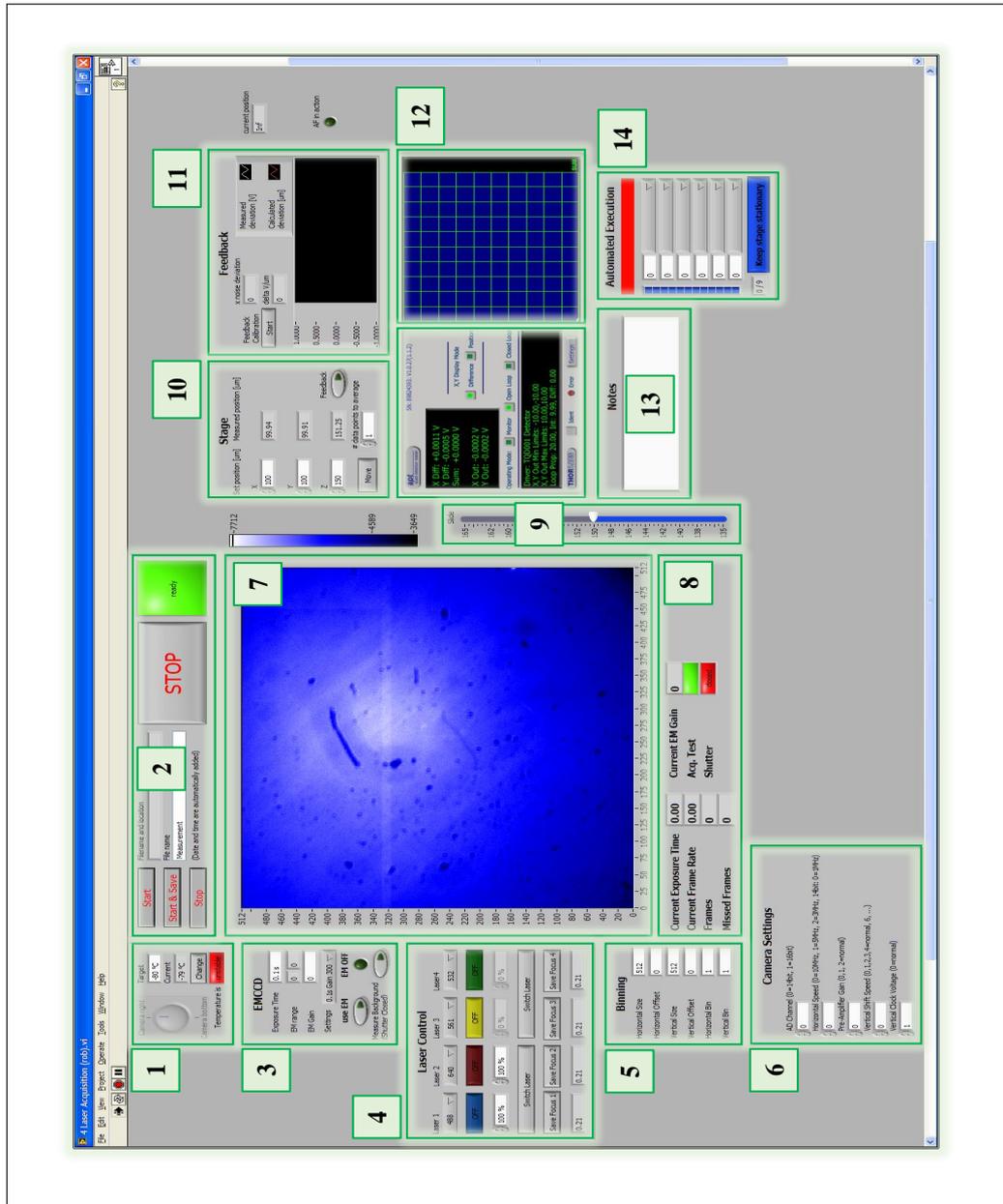


Figure 3.4: The Interface window of the LabView. Green boxes with black numbers were added for annotation. Each part and its explanation were indicated in the text. This figure is adapted from the LabView interface.

### 3.6 Composite Images

---

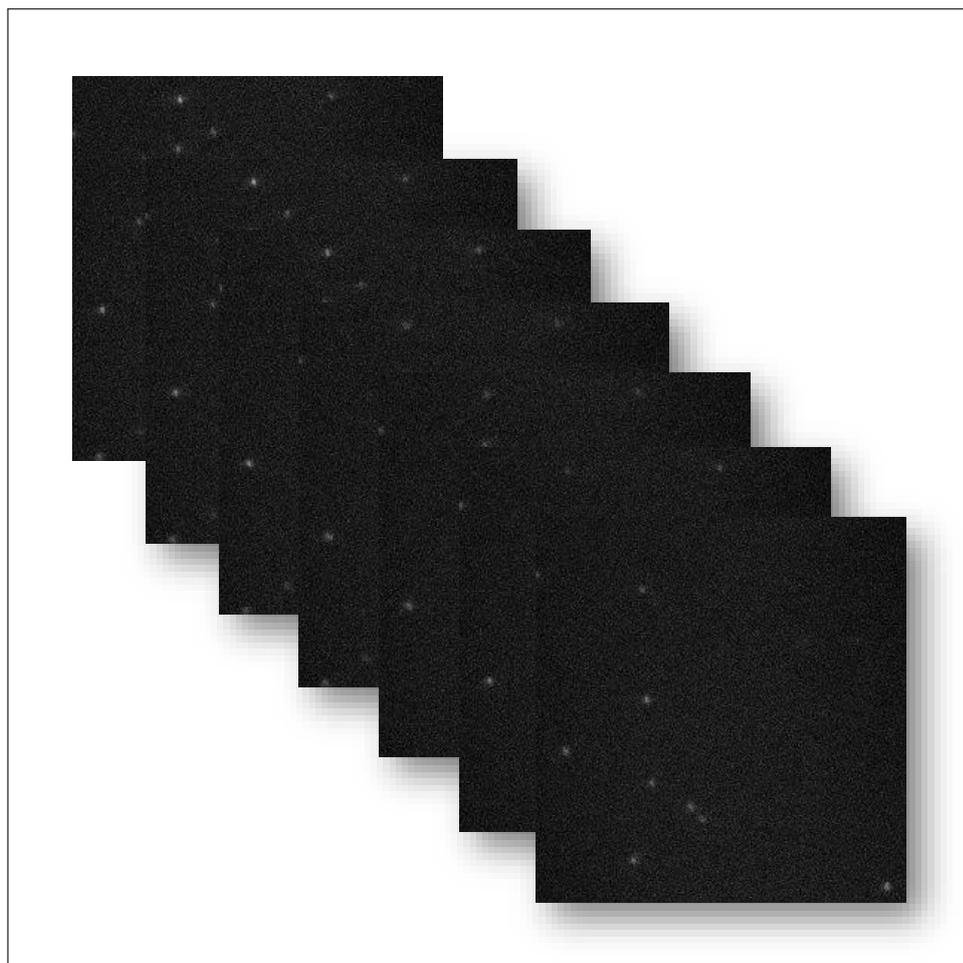


Figure 3.5: Example of a stacked Single Molecule Image. All images were taken by the bottom camera (see figure 3.2). Photos were stacked together by Labview program and saved as they were taken.

### 3.6 Composite Images

---

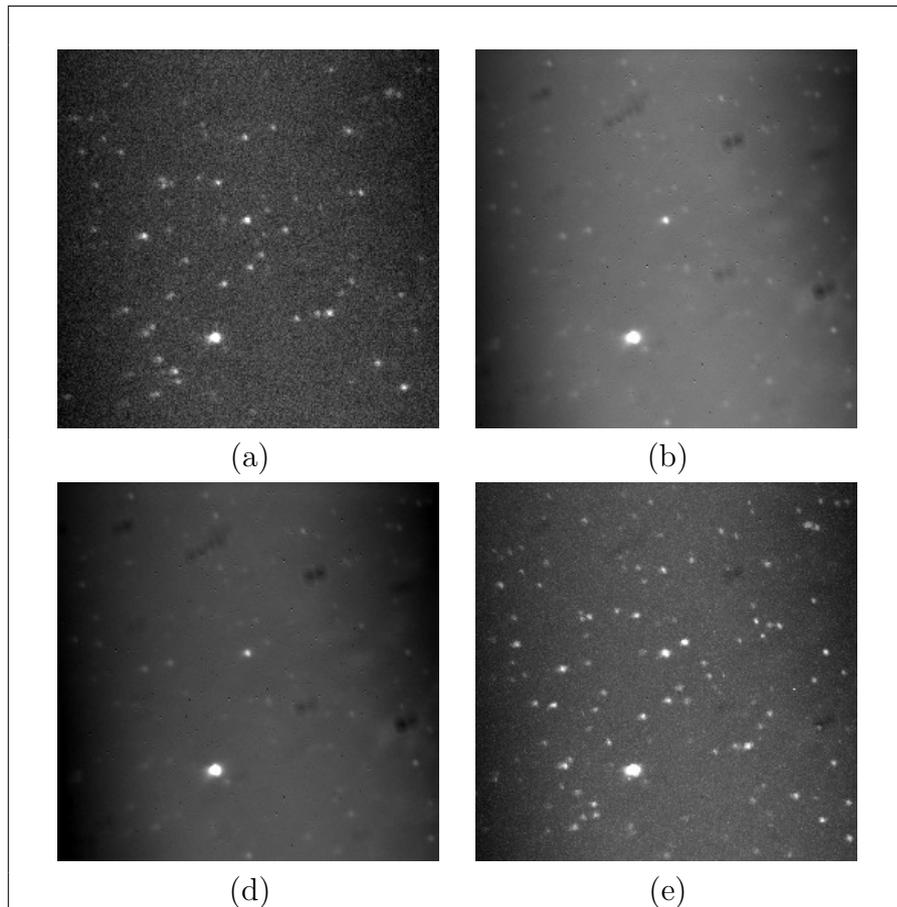


Figure 3.6: Example of a single molecule image images created by calculating the mean or the maximum. All images are from the same acquisition. **(a)** Single image of the fluorescence intensity of GFP molecules under illumination by a wavelength of 488 nm. **(b)** An image created by calculating the SUM value for each pixel over all images with an illumination by a wavelength of 488 nm. **(c)** An image created by calculating the mean value for each pixel over all images with an illumination by a wavelength of 488 nm. **(d)** An image created by calculating the maximum value for each pixel over all images with an illumination by a wavelength of 488 nm. **(e)** An image created by calculating the maximum value for each pixel over all images with an illumination by a wavelength of 488 nm. No colour information is saved in the image.

### 3.7 Spot Detection

---

## 3.7 Spot Detection

Each spot could be identified either by eye or by using an algorithm. The appearance of the composite image was enhanced to be clearer and to help with the identification by eye (Figure3.7 and Figure3.8). The recognition of spots depended on their shape and signal intensity above background. This judgement is subjective, more objective route was to use an algorithm. The one used in this work was written by Dr. Robert Weinmeister [Weinmeister, 2014]. In this research, the algorithm was used.

## 3.8 Inspection of Colocalization

Once the composite images were created, and spots were identified, their colocalization could be easily inspected. Spots of one type of fluorophore were used as a marker. In this study, the marker was Cy5 attached to mRNA. The amount of colocalization with spots of another type of fluorophore, such as a protein attached to GFP, was determined with respect to the marker spots. Each spot of one type of fluorophore can only be colocalized with one spot from another type of fluorophore. However, multiple fluorophores can sometimes be colocalized when appearing as part of one spot due to their distance between each other being smaller than the resolution limit of the microscope. It is probable that there are part of the same complex.

Additionally, colocalization could be determined by eye. The overlap between fluorophores and their colours in the composite images allows pairs of spots to be identified (Figure3.9). This determination can be difficult when the occurring colour shift. A program was implemented in this case to help to establish the colocalization between spots in a faster and objective way.

### 3.8 Inspection of Colocalization

---

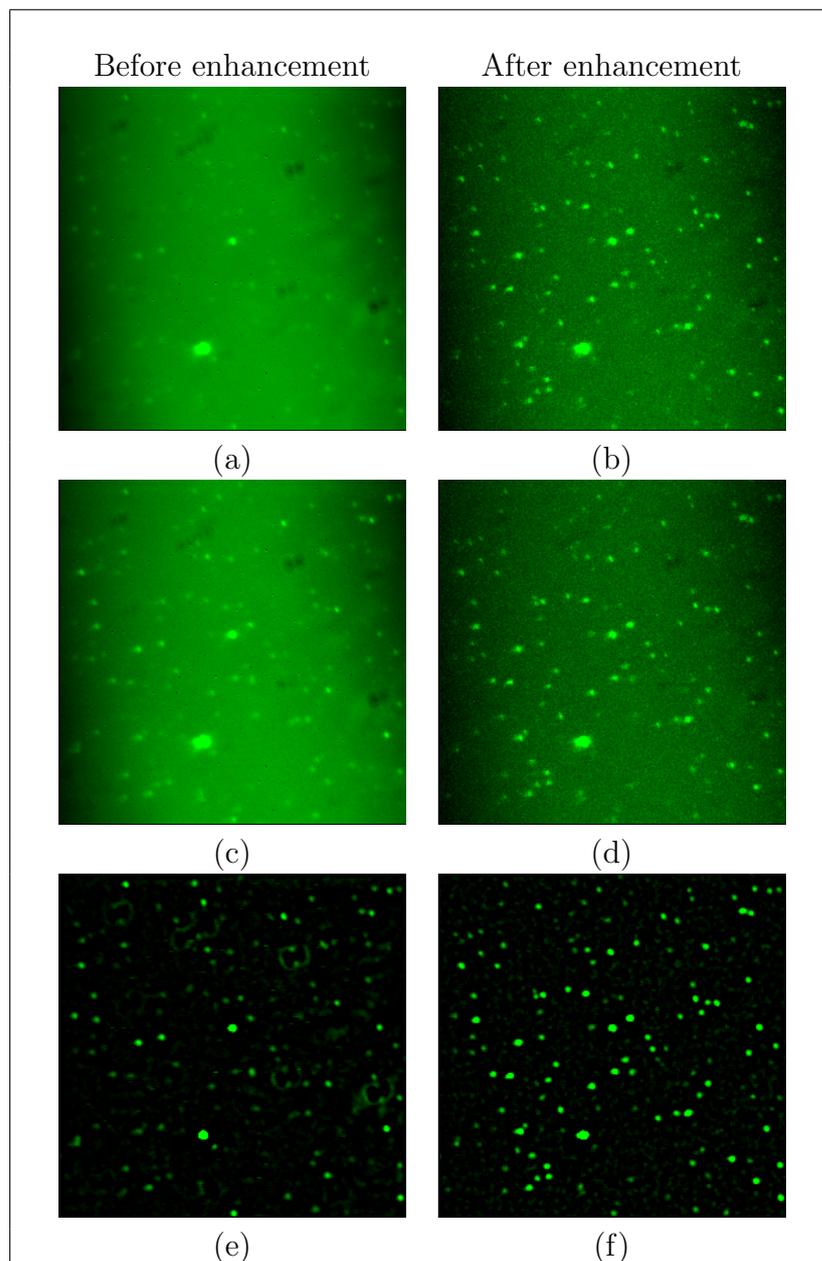


Figure 3.7: Example of a single molecule image before and after the enhancement step. (a) and (b) A coloured image created by calculating the mean value for each pixel over all images with an illumination by a wavelength of 448 nm. (c) and (d) A coloured image created by calculating the SUM value for each pixel over all images with an illumination by a wavelength of 448 nm. (e) and (f) A coloured image created by calculating the MAX value for each pixel over all images with an illumination by a wavelength of 488 nm. All images are from the same acquisition. Right column represents images before any further enhancement. Left column represents the enhanced images.

### 3.8 Inspection of Colocalization

---

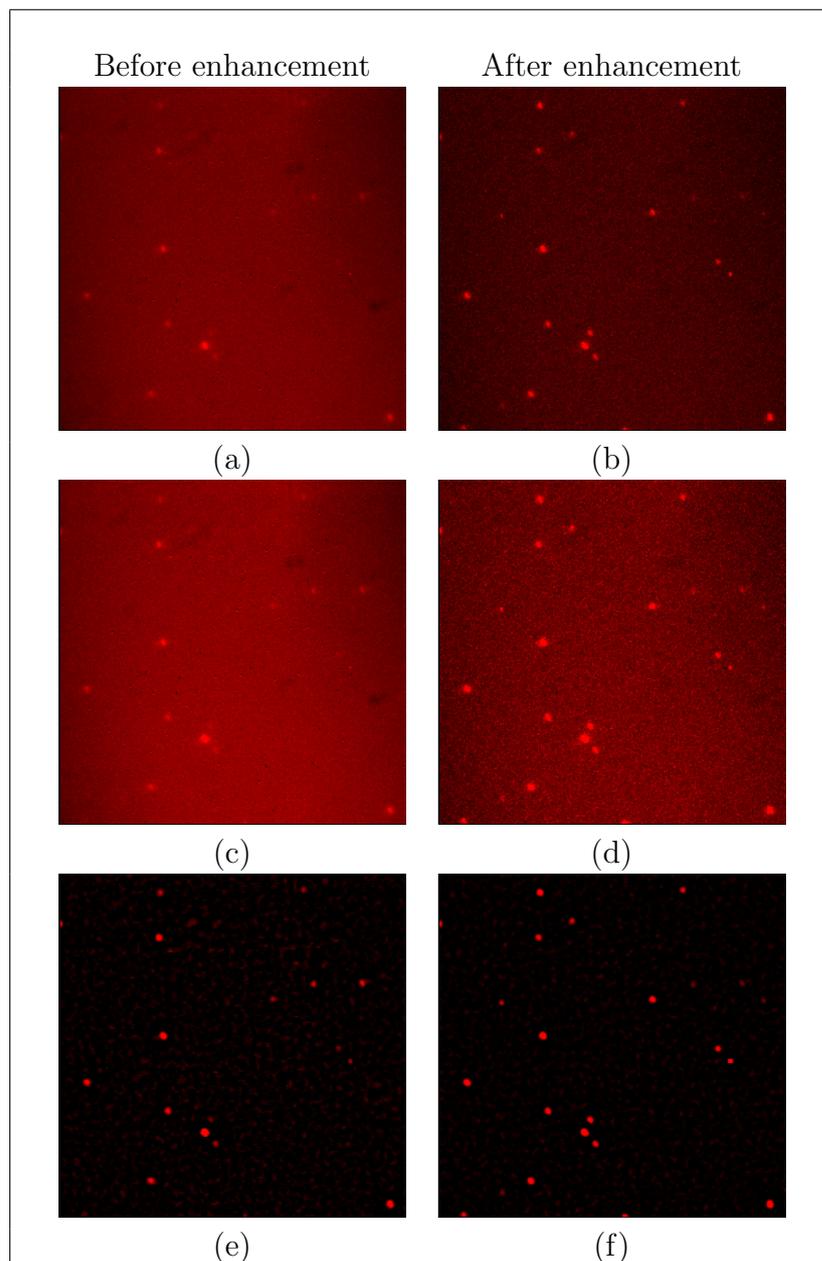


Figure 3.8: Example of a single molecule image before and after the enhancement step. (a) and (b) A coloured image created by calculating the mean value for each pixel over all images with an illumination by a wavelength of 632 nm. (c) and (d) A coloured image created by calculating the SUM value for each pixel over all images with an illumination by a wavelength of 632 nm. (e) and (f) A coloured image created by calculating the MAX value for each pixel over all images with an illumination by a wavelength of 632 nm. All images are from the same acquisition. Right column represents images before any further . Left column represents the enhanced images.

### 3.8 Inspection of Colocalization

---

Further analysis introduced in order to identify the intensity time traces of colocalized spots for bleaching steps to verify the numbers of fluorophores present.

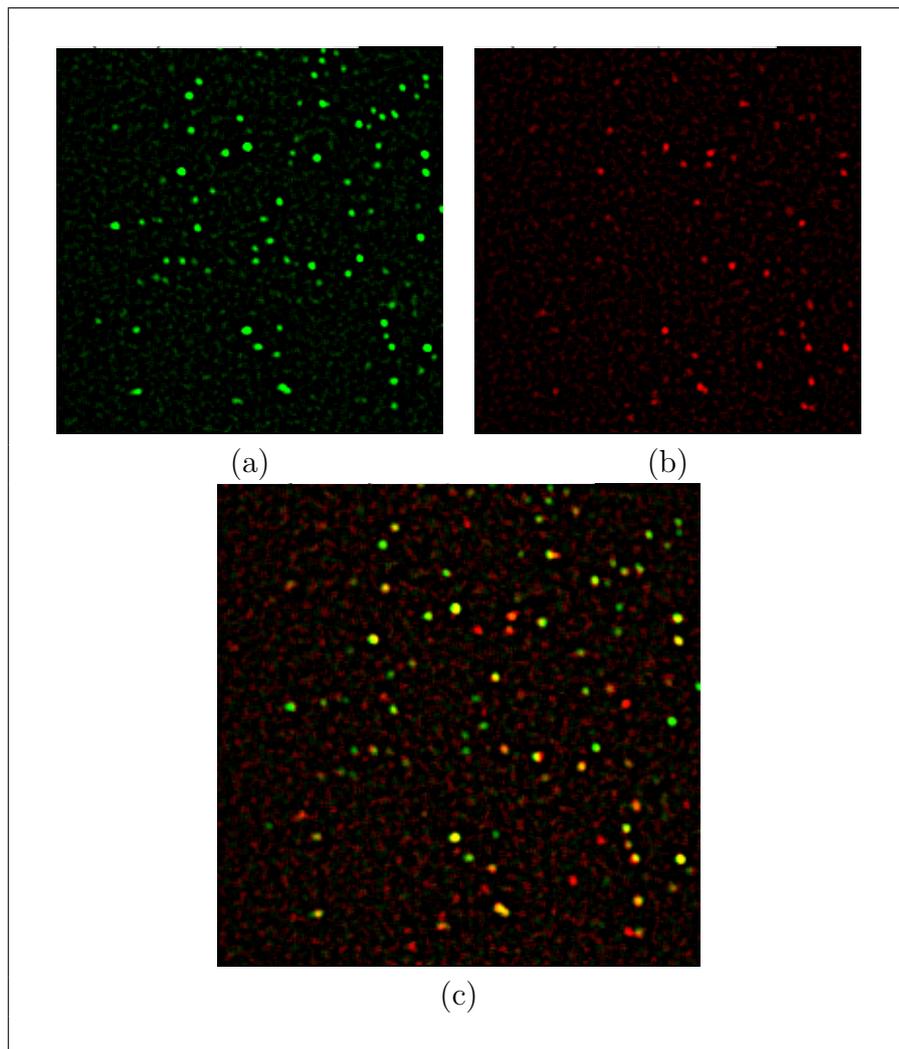


Figure 3.9: Resulted images from the LabView interface. All images are created from the same acquisition. (a) Colour image created by averaging over all images with an illumination by a wavelength of 488 nm, showing spots originating from GFP. (b) Colour image created by averaging over all images with an illumination by a wavelength of 633 nm, showing spots originating from Cy5. (d) The Overlay of the images (a) and (b). Colocalized spots can be seen as yellow.

## 3.9 The Assignment of Steps

As the single splicing complex could not be distinguished directly, traces of intensity vs. time were used in order to assess the number of fluorophores within the complex. Changes of the intensity over time due to the bleaching of fluorophores are important. Counting the steps allowed number of the fluorescent proteins present in the complex to be determined, because each step indicates the bleaching of a fluorophore. These steps were assigned by using an algorithm. The algorithm implemented for this work was written by Dr. Robert Weinmeister [Weinmeister, 2014].

Figure 3.10 shows an example of the different bleaching steps. In the absence of a fluorophore, no step was detected. If there is one protein molecule bound to the RNA, a single bleaching step was seen (Figure 3.10, (a, b and c)). Two bleaching steps indicate the presence of two molecules of proteins bound to the RNA (Figure 3.10, (d)). Three bleaching steps regarding three molecules of proteins bound to RNA (Figure 3.10, (e)). Other cases shows multiple molecules of proteins are bound to the same RNA, this causes multiple bleaching steps which can be observed at a different period of time in different frames (Figure 3.10, (f)).

## 3.10 Analyzing Data

The data obtained by the microscope are sequential intensity images. This data were saved in a tiff file and accompanied by a text file that contained additional information. Composite images were created first. Fluorescent spots had to be identified, and their positions of these spots had to be determined accurately. Then intensity time traces were extracted from the raw data and

### 3.10 Analyzing Data

---

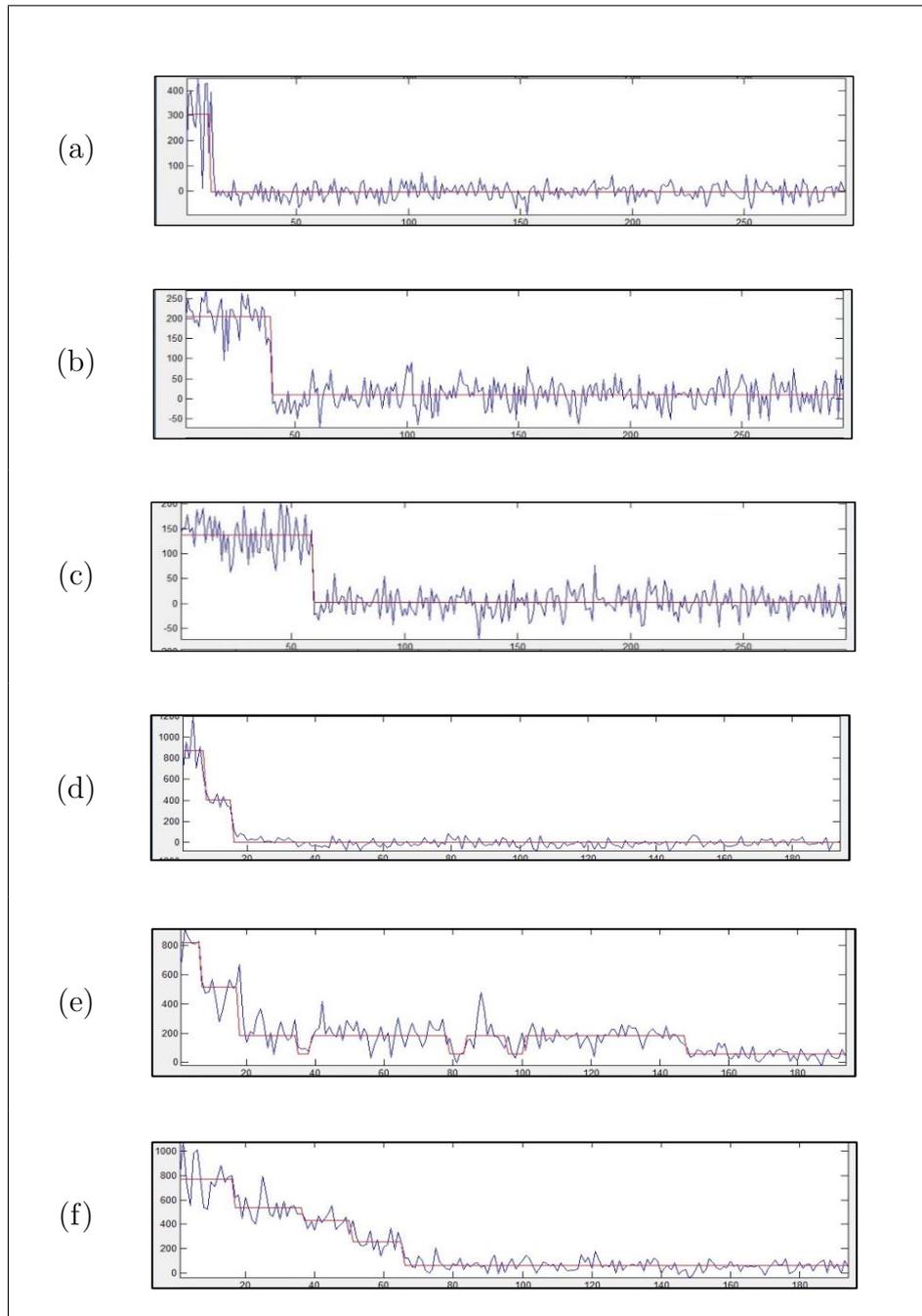


Figure 3.10: A diagram representation of intensity time traces for GFP with a time resolution of 0.1 s. The intensity is shown in blue, the algorithmic fit is shown in red. From (a) to (c) single bleaching step is appeared. (d) Two bleaching steps. (e) Three bleaching steps and (f) Bleaching in multiple steps.

### 3.11 Presentation of Data

---

steps assigned. The number of these steps and the amount of overall colocalisation was recorded. The data were analysed via MatLab program which was written by Dr. Robert Weinmeister [Weinmeister, 2014]. The MatLab allows both simplify the process and reduce the time necessary for it (Figure 3.11).

In this research thesis, the pre-mRNA was labelled with Cy5. This RNA was excited with 633 nm. The spots are represented in red. The mEGFP-labelled proteins were excited with the 488 nm laser which are appeared as green. After the images were stacked, the coloured and enhanced spots were selected either automatically or manually. Once the RNA-Cy5 spots (red spots) overlapped with the GFP (green spots), the central brightest pixel was determined within the MatLab and the colocalised spots are identified. White circles represent the colocalised spots, and the blue ones refer to un-colocalised spots. Also, the enhancement option for the image was available. After the inspection and counting of bleaching steps, the raw data was then exported to an Excel file.

### 3.11 Presentation of Data

The Excel file which exported via Dr. Robert Weinmeister's algorithms [Weinmeister, 2014] represented the percentage efficiency of the protein binding for each bleaching step. However, to compute the number of molecules regarding each step, calculation has to be done manually for each Excel file. Furthermore, the calculation of the error bars has to be done manually too. Moreover, each resulted bar chart has to be extracted manually in order to be represented in a document. Those additional manual processes are time consuming and might lead to an unintended errors.

### 3.11 Presentation of Data

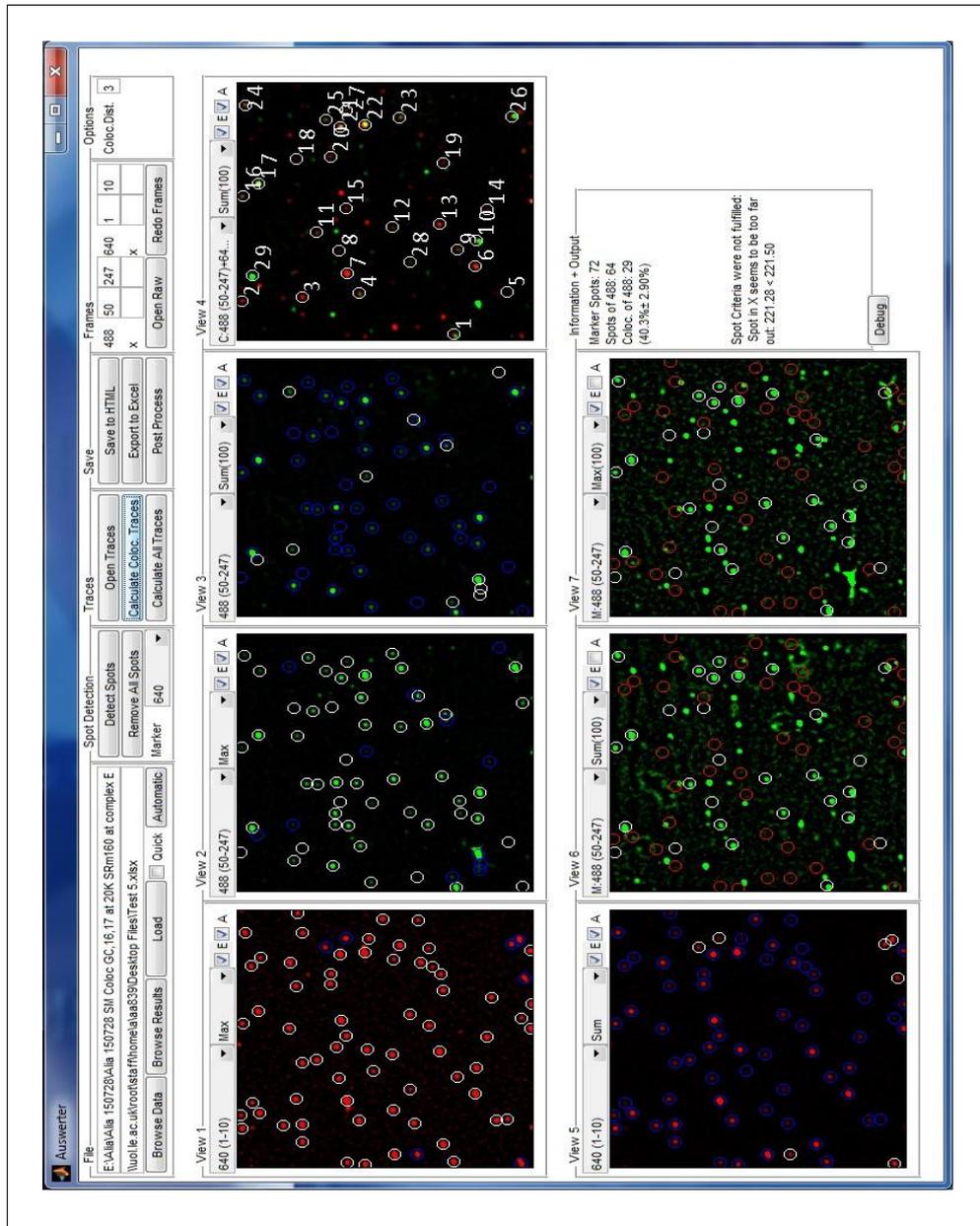


Figure 3.11: Schematic diagram of the interface of the MATLAB Program. MatLab was used in this research in order to detect single spots and to analyse data. This program was developed by Dr. Robert Weinmeister [Weinmeister, 2014].

### 3.11 Presentation of Data

---

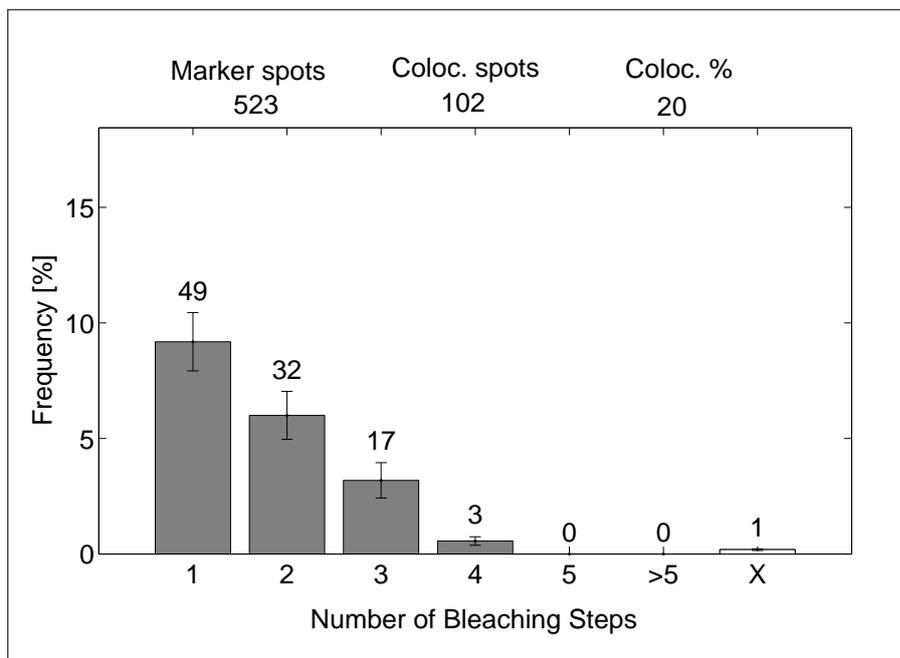


Figure 3.12: Exemplary representation of single molecule colocalisation experiment result. The diagram represents an experiment With two colours, one for the marker fluorophores (Cy5) and one for the colocalized fluorophores (GFP). From the top left, marker spots represents the number of marker fluorophores spots (Cy5). Colocolisation spots represents the number of GFP spots identified in the experiment. Coloc. % represents the percentage of the colocalised spots.

Due to these limitations, I developed a tool by Matlab with help from Dr. Alaa Khadidos to overcome these limitations. The tool reads the corresponding cells from each Excel file automatically to calculate the number of molecules regarding each step and to compute the overall percentages of the colocalisation efficiency, number protein molecules and error bars. Automatically all resulted data are stored in different formats such as .EPS and .PDF to easily represented in a document.

The results of the single molecule experiments are presented as bar charts (Figure 3.12). The bar chart shows the distribution of the number of steps found in the intensity time traces of the colocalised spots.

### 3.11 Presentation of Data

---

On top of the bar I show the total number of identified marker spots, the number of spots colocalised with these spots and the resulting colocalisation percentage. The white box on the right represents the accumulated percentage for all spots that had more identified steps or for which no certain number of bleaching steps have been identified. This could be due to a fluorophore that did not bleach or where the trace did not allow a clear assignment of the number of bleaching steps than individually shown on the grey bars to the left.

## Chapter 4

# Optimisation of Condition for Single Molecule experiments with SRm160

Multiple steps have to be taken to prepare for the eventual experiments. A functional home-made nuclear extract was prepared from cells expressing modified proteins. Both the protein of interest and pre-mRNA constructs have to be labelled with fluorescent markers in order to make them visible for fluorescence microscopy. The molecules of interest were confined to the surface of the sample chamber to allow investigation over time at a distinct splicing complex.

# Contents

4.1	Labelling of SRm160 in Nuclear Extract . . . . .	116
4.2	Labelling of the Pre-mRNA . . . . .	121
4.3	Tethering the Surface and Modifications . . . . .	121
4.4	Preparation of Splicing Complexes . . . . .	123
4.4.1	Stalling Splicing at the Early Complex E . . . . .	123
4.4.2	Stalling Splicing at Complex A . . . . .	124
4.4.3	Stalling Splicing at Complex C . . . . .	124
4.5	Pre-mRNA Concentration Required for Single Molecule Experiments . . . . .	125

## 4.1 Labelling of SRm160 in Nuclear Extract

To investigate how SRm160 protein binds to a pre-mRNA substrate via TIRF microscopy, the first step was to label the protein of interest (here is SRm160 protein) in nuclear extracts with a fluorescence component. The SRm160 plasmid used in this research was a kind gift from Dr. Jeffrey Nickerson. The protein SRm160 was labelled by expressing it as a recombinant protein fused to mEGFP (Chapter 2, section 2.2). Subsequently, SRm160 plasmid was transfected into 293T cells to allow expression of recombinant proteins. Expression was then detected by fluorescence microscopy (Figure 4.1). After that, cells were harvested for the further preparation of nuclear extracts.

The prepared nuclear extract was shown to be functional in *in vitro* splicing among with GloC pre-mRNA. This contains two exons and an intron, derived from a rabbit  $\beta$ -globin gene. The intron and second exon were truncated, and the 5' splice site was a consensus sequence [Hodson et al., 2012]. Figure 4.2 shows that the extract was active and ready to implement into further experiments. Moreover, the activity of this extract was also tested in other constructs such as GloC-E3 (no  $\alpha U1$  binding site at the 3' end exists), GloC-SMN2 and GloC-SMN2+ESE, see Figure 4.3.

Additionally, fluorescent westerns blots (Figure 4.4) were done on the nuclear extract to assess the proportion of GFP-SRm160 in the extract that was labelled and the concentration compared to that in a normal extract. The mEGFP fusions combined 90% of the SRm160 in the extract. All single molecule colocalisation experiments in this thesis work were performed with the prepared GFP-SRm160.

#### 4.1 Labelling of SRm160 in Nuclear Extract

---

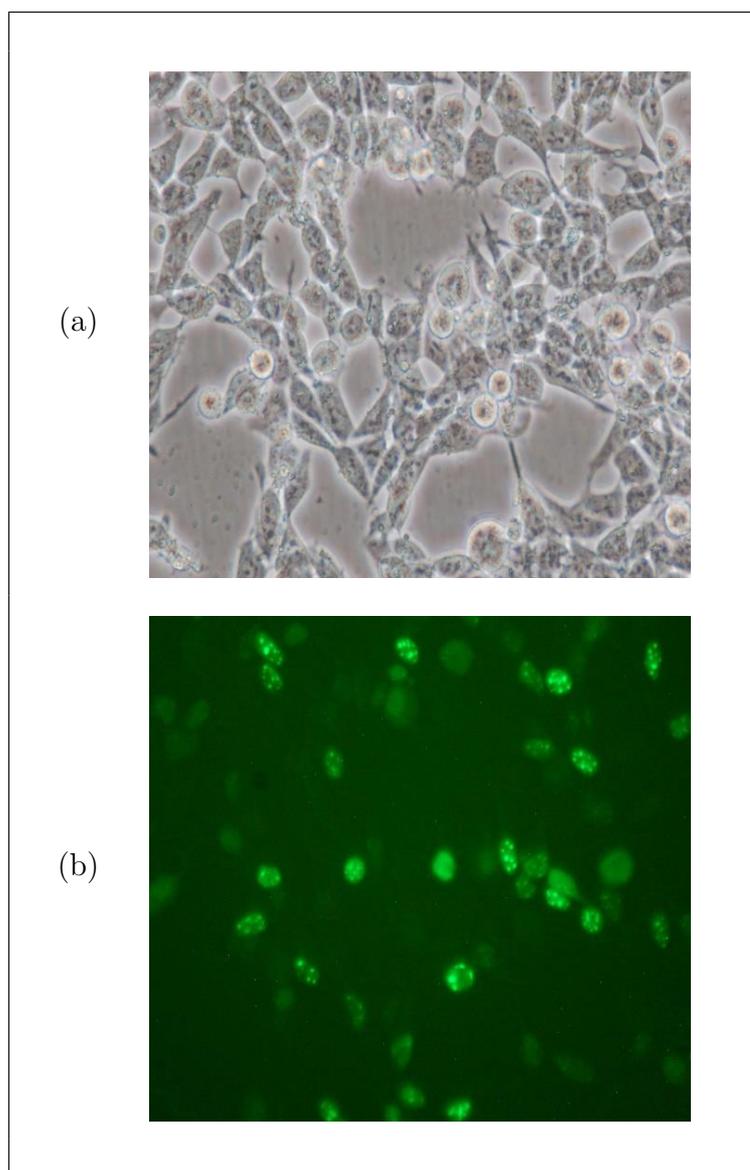


Figure 4.1: Fluorescence microscopy view of expressed protein. GFP-SRm160 was transfected into 293T cells. (a) represents white image of the transfected protein into 293T cells. (b) represents the fluorescent protein (in green) transfected into 293T cells.

## 4.1 Labelling of SRm160 in Nuclear Extract

---

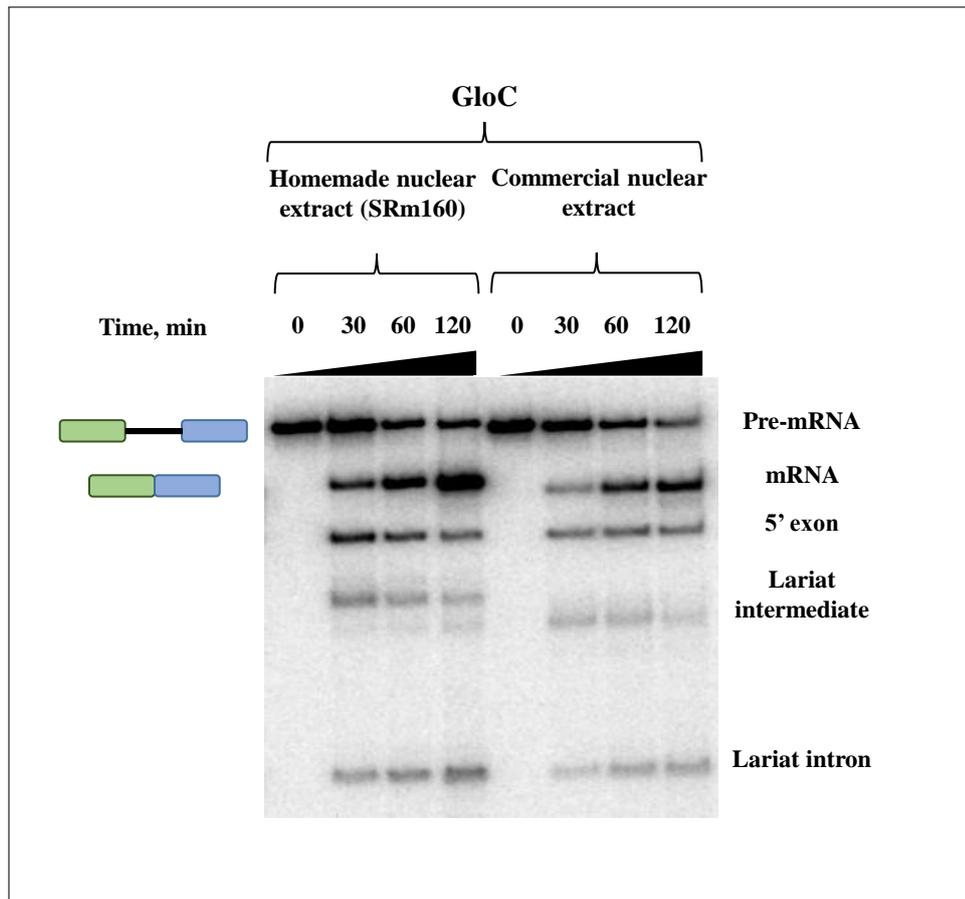


Figure 4.2: Control *in vitro* splicing reaction. Splicing experiments confirm the functionality of the home-made nuclear extract (SRm160) with GloC construct. A 6 % polyacrylamide gel was used.

## 4.1 Labelling of SRm160 in Nuclear Extract

---

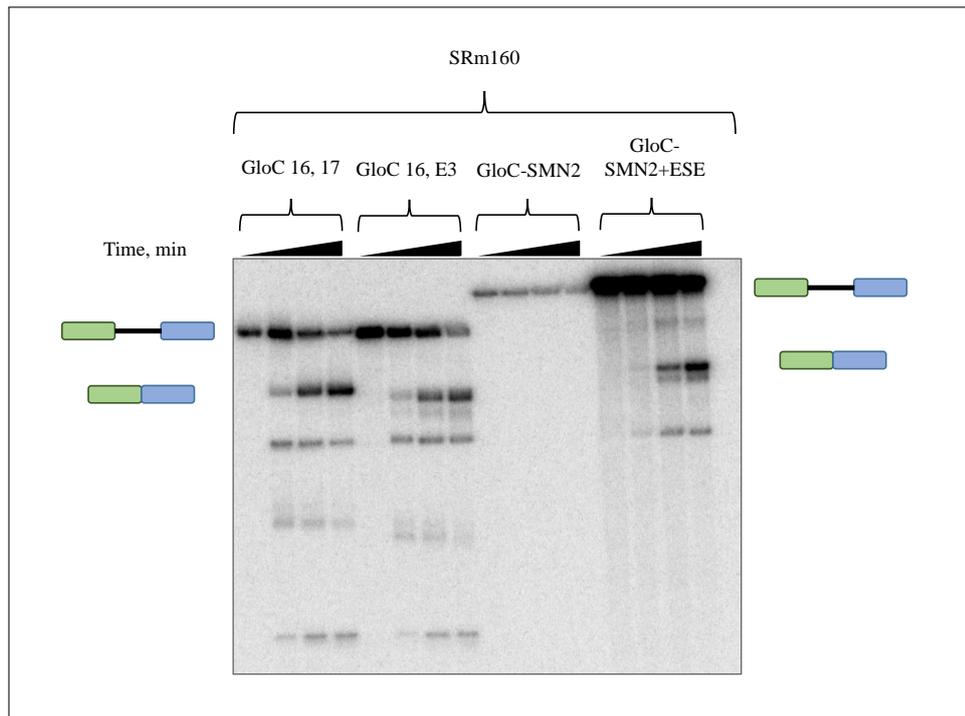


Figure 4.3: Control *in vitro* splicing reaction. Splicing experiments confirm the functionality of the home-made nuclear extract (SRm160) in different construct. From left; GloC 16, 17 refers to GloC substrate with extra U1 binding site at the 3' end (the numbers 16 and 17 are represent the primers name used in the experiment), GloC 16, E3 refers to GloC substrate without extra U1 binding site at the 3' end (E3 represents the primer's name used in the experiment), GloC-SMN2 refers to a substrate that lacks the ESE binding site (for more details see Chapter 5, section 5.1.2) and the GloC-SMN2+ESE refers to a substrate that does not lack the ESE binding site (for more details see Chapter 5, section 5.1.2). A 6 % polyacrylamide gel was used.

#### 4.1 Labelling of SRm160 in Nuclear Extract

---

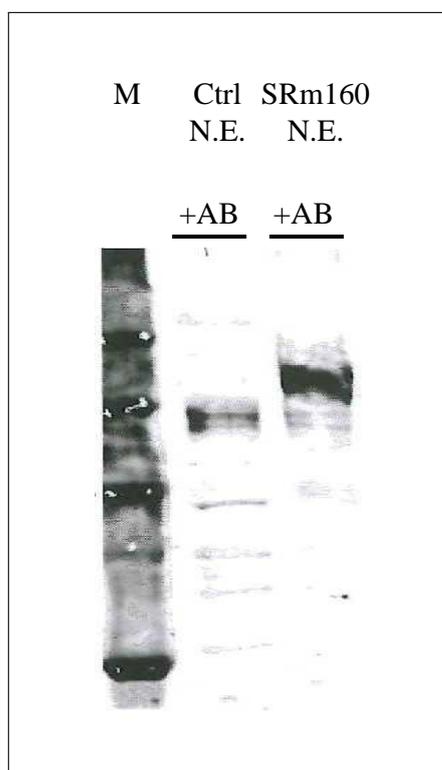


Figure 4.4: Control experiment of western blot. Fluorescent western blot of the GFP-SRm160 nuclear extract and the commercial nuclear extract. Western experiments was probed with an antibody specific for SRm160. From left, (M) represents the protein marker, (Ctrl N.E. +AB) represents the commercial nuclear extract in the presence of suitable antibody and (SRm160 N.E. +AB) represents the SRm160 protein in the presence of an antibody specific for SRm160.

## 4.2 Labelling of the Pre-mRNA

A fluorescent fluorophore has to be attached to a single molecule of pre-mRNA. The oligonucleotide used in this research work is  $\beta$ -globin-5'-Cy5. This oligo has Cy5 at one end and a biotin group at the other end (8.1.23). Once the oligo is annealed to its complementary sequence, the substrate is ready to use for single molecule purposes. It is therefore crucial to ensure that the oligo used for labelling the pre-mRNA is annealed quantitatively such that no free oligos remain. This can be achieved by optimising the annealing conditions and using an excess of pre-mRNA compared to the amount of oligo (Figure 4.5).

The annealing of the labelled-oligo (Cy5-oligo) to the pre-mRNA was done with increasing amounts of pre-mRNA (see Chapter 2, section 2.3.1). The RNA was run on a 6 % non-denaturing polyacrylamide gel and its fluorescence scanned via Typhoon (Figure 4.5). The amount of free Cy5-oligo was assessed. It decreases with increased amount of pre-mRNA. Higher amounts of pre-mRNA do not allow for less free Cy5.

## 4.3 Tethering the Surface and Modifications

As the fluorescent probes have to be close to the surface for observation under total internal reflection (TIR) illumination, a commonly used system of biotin and streptavidin was employed and evaluated by Dr. Robert Weinmeister [Weinmeister, 2014]. The surface was incubated with BSA-biotin and streptavidin. The BSA-biotin binds unspecifically to the surface. The streptavidin binds to the biotin attached to the BSA on the surface and also provides further available binding sites for the biotin of the labelling oligo as it has four binding

### 4.3 Tethering the Surface and Modifications

---

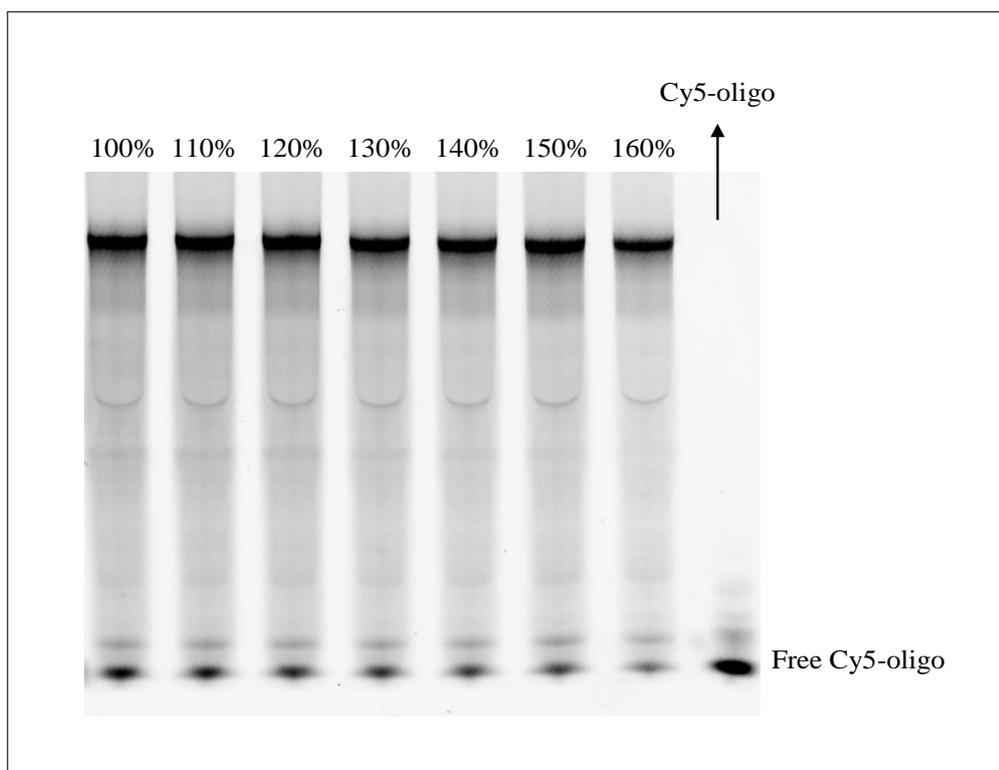


Figure 4.5: Labelling of pre-mRNA via 5'-Cy5-oligonucleotide. Increasing amounts of pre-mRNA compared to the labelling oligo are run on a 6 % denaturing polyacrylamide gel and its fluorescence imaged. From left; 100 % to 160 % are different percentages of pre-mRNA molecules to molecules of the labelling oligo, e.g., 100 % means the same number of pre-mRNA molecules as molecules of the labelling oligo. The Cy5-oligo at the very right of the figure represents a sample of Cy5-oligo only (no pre-mRNA was added).

## 4.4 Preparation of Splicing Complexes

---

sites for the biotin. These interactions allow the pre-mRNA to be confined to the surface. The amounts of streptavidin and BSA-biotin to pre-mRNA and protein spots were also optimized. Different amounts of streptavidin and BSA-biotin were used. A normal splicing reaction was also observed. Neither the concentration of BSA-biotin nor the concentration of streptavidin influenced the amount of pre-mRNA spots. Pre-mRNA was found on the surface even in the absence of the BSA-biotin and streptavidin. As a conclusion of Dr. Robert Weinmeister experiments, it was decided to not use the streptavidin-biotin system [Weinmeister, 2014]. Additionally, the splicing reactions were diluted to allow analysis of single molecules. Other modifications of the surface were tested in order to reduce the unspecific binding. None of these modifications showed any sufficient differences such that no additional modifications of the surface were used [Weinmeister, 2014].

## 4.4 Preparation of Splicing Complexes

The next step after the preparation of nuclear extract of interest is to look at protein binding in different splicing complexes. This process can be implemented by stalling the splicing reaction at different complexes such as E, A and C.

### 4.4.1 Stalling Splicing at the Early Complex E

Splicing is a process that is characterized by the formation of several complexes. Such complexes are the early complex E, complex A, and complex C. The splicing reaction can be stalled to obtain homogeneous samples representing a chosen complex in order to make the investigation of each state applicable.

## 4.4 Preparation of Splicing Complexes

---

The first complex that forms during splicing is the early complex (E). This complex has only a U1 snRNP bound to it at the 5' splice site (5' ss) along with some other splicing regulator proteins (factors). The depletion of ATP stalls splicing at the early complex (E) [Michaud and Reed, 1991]. Stalling splicing at the early complex E was confirmed by native gel electrophoresis [Li, 2016].

### 4.4.2 Stalling Splicing at Complex A

The next complex to form after the early complex (E) is complex (A). This complex is formed by the replacement of SF1 factor with U2 snRNP at the branch point [Hastings and Krainer, 2001] in the presence of ATP. As mentioned previously, the splicing reaction can be stalled to obtain homogeneous samples representing a chosen complex. Stalling the splicing reaction at complex A can be achieved using various reagents. A successful stalling of the splicing reaction at complex A were achieved and compared by Dr. Mark Hodson [Hodson, 2011]. In this thesis, it was decided to implement anti-U6-oligonucleotide ( $\alpha$ U6), a modified RNA oligonucleotide, to prevent U6 snRNA base pairing with U2 snRNA [Dönmez et al., 2004]. Native gel electrophoresis showed that the addition of  $\alpha$ -U6 oligo successfully blocked the formation of the splicing complex A. Therefore,  $\alpha$ -U6 oligo has chosen to be the reagent to promote complex A formation [Li, 2016].

### 4.4.3 Stalling Splicing at Complex C

The next complex in splicing reaction is complex C. Blockage at complex C is quite simple. The conserved 3' splice site (3'ss) consists of AG dinucleotide and simple mutation at this site leads to the formation of complex C but

#### 4.5 Pre-mRNA Concentration Required for Single Molecule Experiments

---

the pre-mRNA does not splice. Therefore, modifying the 3' splice site of the pre-mRNA (Glo C GG) leads to accumulation of complex C [Jurica et al., 2002]. Stalling the splicing gel at complex C was confirmed by native gel electrophoresis [Li, 2016].

### 4.5 Pre-mRNA Concentration Required for Single Molecule Experiments

In single molecule experiments, the concentration of pre-mRNA used was 62.5 nM. The pre-mRNAs used in these single molecule experiments are labelled by 5'-Cy5 oligo. It is therefore important to know whether this concentration and the labelling of pre-mRNA affects the efficiency of splicing and complex formation.

Firstly, the splicing of GloC was tested at various concentrations; 62 nM and 31 nM in a single molecule reaction. It was shown both <sup>32</sup>P-labelled GloC and Cy5-labelled GloC were spliced *in vitro*.

Also, it was necessary to test whether all complexes formed properly with this concentration. Native gels were run with GloC construct using no ATP condition (at the early complex E) and with  $\alpha$ U6 oligo (at complex A). The complexes were seen by <sup>32</sup>P labelled uncapped GloC at very low concentration with different concentrations of unlabelled GloC in the same reaction. The formation of both E and A complexes on GloC did not change upon addition of high concentration of unlabeled transcribed Globin C.

Moreover, Western blot experiments demonstrated that the protein concentrations are relatively high in the nuclear extracts; they are normally several  $\mu$ M, which is much higher than nM. This indicates that pre-mRNA concen-

## 4.5 Pre-mRNA Concentration Required for Single Molecule Experiments

---

tration of 62.5  $\mu\text{M}$  in a typical single molecule experiment is acceptable to use.

Additionally, the formation of complex B and C by 62.5 nM Cy5-cap-labelled GloC were also tested by native gel. It was shown that all of the pre-mRNA was stuck at complex B at 30 min time point in the addition of a concentration of 62.5 nM of GloC pre-mRNA.

Therefore, to allow complex C to form, a concentration gradient was done using a range of different concentrations of unlabeled GloC pre-mRNA with a trace  $^{32}\text{P}$ -labelled GloC as an indicator. Analysis of the native gel showed that 15 nM GloC is the highest GloC concentration to form Complex C in the standard single molecule reaction [Li, 2016].

Therefore, in subsequent single molecule experiments the pre-mRNA concentrations used were 62.5 nM for E, A and B complexes and 15 nM for complex C.

## Chapter 5

# Behaviour of SRm160 protein during splicing.

*So far, it is not clear how many molecules of SRm160 protein are bind to the pre-mRNA construct during the splicing reaction. Here in this thesis work, the binding of SRm160 protein during the formation of splicing complexes; the early complex E and complex A are investigated. How the number of bound SRm160 proteins is influenced by the presence or the absence of other factors such as the U1 snRNP, U2 snRNP, phosphorylation (PhoStop), an ESE sequence sites and Tra2 $\beta$  sites are also investigated.*

# Contents

5.1	Introduction . . . . .	129
5.1.1	The State of Dimerization . . . . .	130
5.1.2	The Construct . . . . .	131
5.1.3	Control Experiments . . . . .	132
5.2	SRm160 in Different Complexes; <i>the scope from different constructs</i> . . . . .	138
5.2.1	GloC Construct . . . . .	138
5.2.1.1	GloC (Non-Modified Construct) at the Early Complex E (the depletion of ATP) . . . . .	138
5.2.1.2	Globin C Construct at Complex A (the addition of ATP) . . . . .	140
5.2.2	The Interplay of SRm160 Protein With the U1 snRNP and the 5' Splice Site; Globin M . . . . .	143
5.2.3	Intron-less Globin C; mRNA . . . . .	144
5.3	SRm160 and the Splicing Enhancer; SMN2 . . . . .	146
5.4	SRm160 and the Tra2 Site . . . . .	154
5.5	SRm160 Vs. the 5' exon . . . . .	161

### 5.1 Introduction

So far, there is no research study that identified the role of SRm160 splicing coactivator protein (consists of the serine/arginine (SR)-related nuclear matrix protein of 160 kDa) at each distinct splicing complexes. Less than 30 research papers were published. Their investigations range from the localisation of SRm160 protein in the cell to the relation between this protein and the splicing products (i.e., lariat intermediate and lariat product). Additionally, there is no clear studies specific for a distinct splicing complexes; such as the early complex E or complex A or complex C. Because of these limitations, the idea of this research work has born.

Using single molecule techniques allow the role of splicing proteins and other factors in recognition and selection of the splice sites to be easily identified. This chapter shows that SRm160 protein associates efficiently with the presence of U2 snRNP, but not with U1 snRNP, at condition supports the formation of the early complex E. It's binding also depends on the presence of an ESE sequence site. It has been suggested that the SRm160 protein might recognise candidate sites and might be involved in the early splicing complex E and complex A formation and hence propose a restructuring of the pre-mRNA complexes.

In order to identify the state of SRm160 protein binding to a pre-mRNA at a distinct splicing complex; splicing experiments were performed at single molecule level. Such that, few experiments have to be identified first before any further single molecule experiments in order to ensure the accuracy of the obtained results. The next following sections include the state of dimerisation of SRm160 protein, all constructs used in this work and other

## 5.1 Introduction

---

control experiments.

### 5.1.1 The State of Dimerization

In a splicing reaction, number of proteins can interact with each other and hence multimerise to mediate splicing. Therefore, it is important to investigate the state of dimerisation of the protein of interest (SRm160) in the splicing reaction before any further experiments. In this work, snRNAs and any other RNAs left over were digested. This process was performed via the addition of RNase A (RNA digesting enzyme; Ribonuclease) to the splicing reaction.

Figure (5.1) shows the distribution of SRm160 molecules bound to pre-mRNA (GloC); X axis. The Y axis shows the percentage of the binding of SRm160. The total number of pre-mRNA and the number of SRm160 molecules with different bleaching steps are marked on top of the bar. The obtained results showed that SRm160 proteins have different levels of dimerisation. The depletion of ATP (a condition that allows the formation of the early complex E) shows almost one molecule of SRm160 is bound to the GloC pre-mRNA. This experiment results in molecules of SRm160 are bind to GloC with 75%. This result indicates that the SRm160 molecule does not bind as a dimer. However, the level of SRm160 binding is changing up to two molecules bound the pre-mRNA in the presence of ATP and  $\alpha$ U6 oligo (the condition that allows the formation of complex A). One-third of the total number of SRm160 molecules were appeared as a monomer in the absence of ATP while two-third of them were exist as a dimer in the presence of ATP. As a conclusion, addition of RNase A means no RNA molecules are exist and thus the SRm160 molecules are the only residues or supposed to be the only. This hypothesis supposed to be correct because when I test the uncolocalized molecules of

## 5.1 Introduction

---

SRm160 - SRm160 molecules that are not colocalized with the RNA substrate - one significant bar is analysed in the distribution chart. This means that the free SRm160 molecules appear as a single molecule. On the other hand, the two-third of the total number of SRm160 molecules might be an of the presence of the RNase A. By this stage, the analysis of further results and bar charts is become easy to explain and compare.

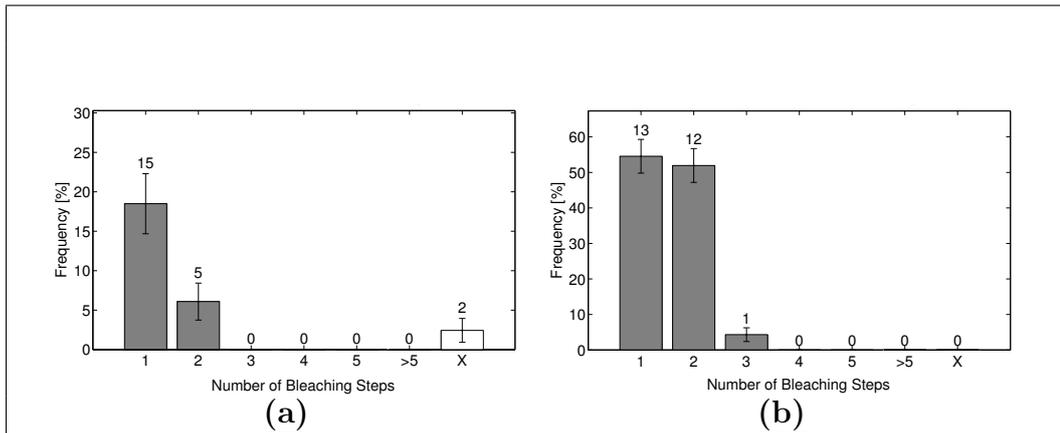


Figure 5.1: The state of the SRm160 protein dimerization in the presence of RNaseA. Splicing reaction at single molecule level were performed at different conditions. **(a)** 22 molecules of SRm160 were identified at the early complex E condition (no ATP). **(b)** 22 molecules of SRm160 were identified at complex A condition (with ATP) and  $\alpha$ U6 oligo. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding.

### 5.1.2 The Construct

To investigate how and how many SRm160 molecule is react with the pre-mRNA during splicing reaction, multiple RNA constructs were used in order to investigate different conditions. Those constructs are GloC, Glo M, and Intronless RNA (mRNA).

Moreover, constructs of SMN2 with and without ESEs sequences were

## 5.1 Introduction

---

also used. Those constructs ,were prepared by Andy Jobbins, provide an opportunity to look at the correlation between simple molecule analysis and a wealth of ensemble studies on the behaviour of SRm160 proteins, U1 snRNP and U2 snRNP on exon 7. Figure (5.2) shows each construct used in this work in the single molecule experiments.

### 5.1.3 Control Experiments

Splicing reaction is a complicated system influenced by multiple factors. The binding distributions could be affected by those factors. Several control experiments were performed to ensure the validity of our findings.

The slide chamber was tested at the beginning of each experiment in order to ensure the validity of the slide. The experiment was performed by looking at the slide under the microscope in the absence of any reagents/buffers. The obtained results in (Figure 5.3) shows no existence of fluorescent spots which might interfere to the real results. This ensures that the slide chamber is applicable to use for single molecule experiments.

The state of GFP was tested under the microscope in order to ensure that there is no signals can contribute to the real results. The experiment was performed using SRm160 protein by itself. No further conditions were implemented. The obtained results show single bleaching step, see (figure5.4).

The Cy5-oligo used to label all constructs was tested too in order to ensure if there is any binding could occur between SRm160 protein and this oligo. The experiment was performed in the presence of SRm160 extract and the Cy-5 oligo. The obtained results indicates that SRm160 to protein does not bind to this oligo. The results ensure that the presence of Cy5-oligo does not contribute the state of SRm160 binding to pre-mRNA construct.

## 5.1 Introduction

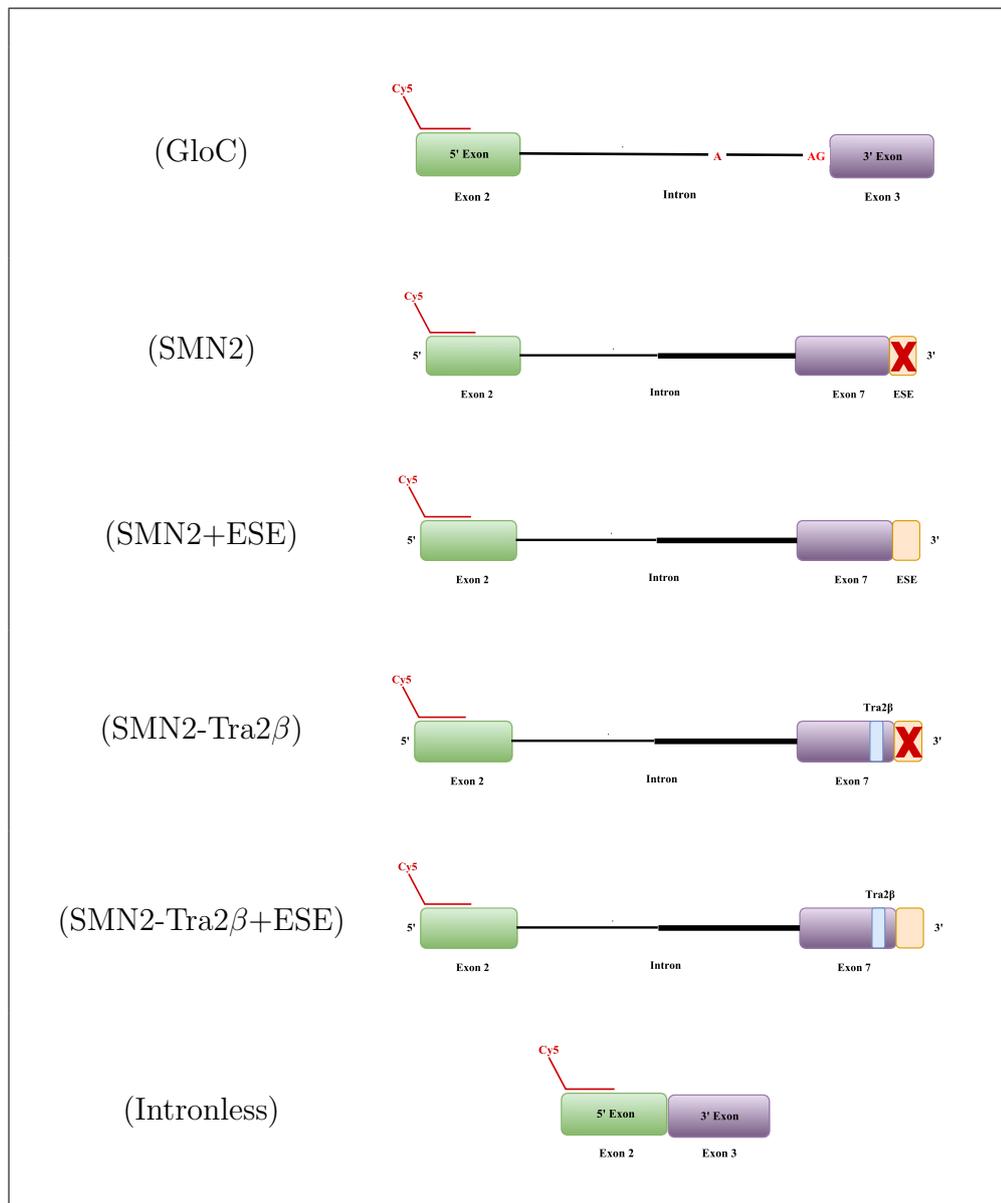


Figure 5.2: Schematic diagram of different constructs of RNA used in single molecule experiment. Each construct is labelled by its name on the left side of the figure. From left; green box represents 5' exon, solid black line represents the intron and purple box represents 3' exon. Solid red line indicates the Cy5-labelled oligonucleotide.

## 5.1 Introduction

---

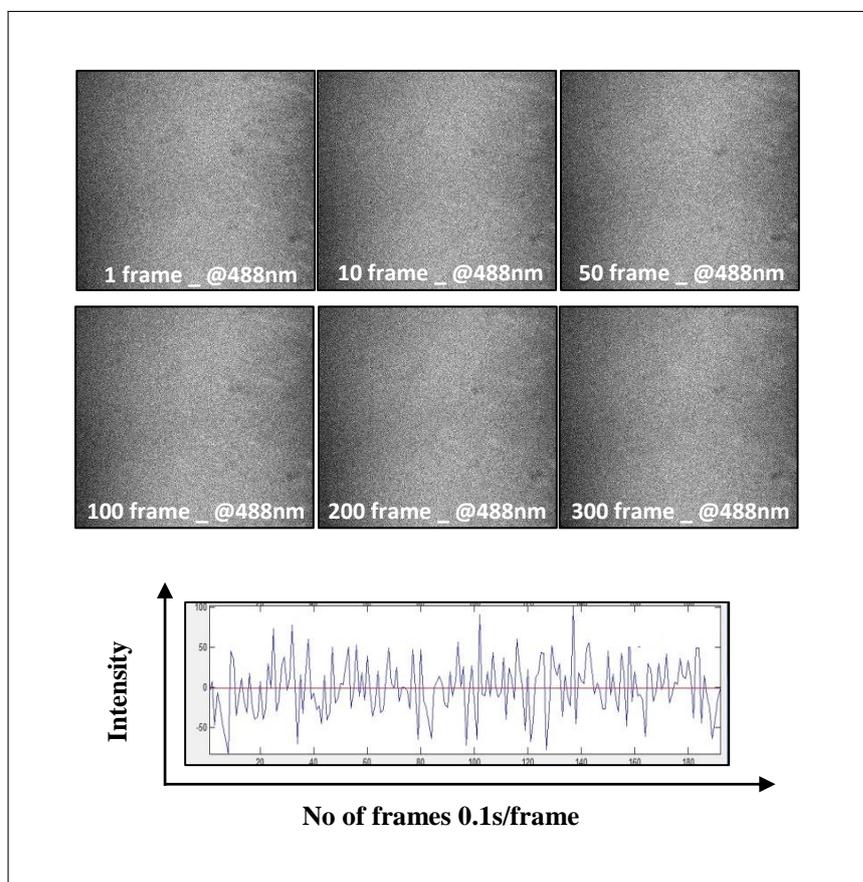


Figure 5.3: Testing the purity of the slide chamber for single molecule colocalisation experiments. The two top rows are series of detected images of the slide chamber under the microscope. No spots are detected.

## 5.1 Introduction

---

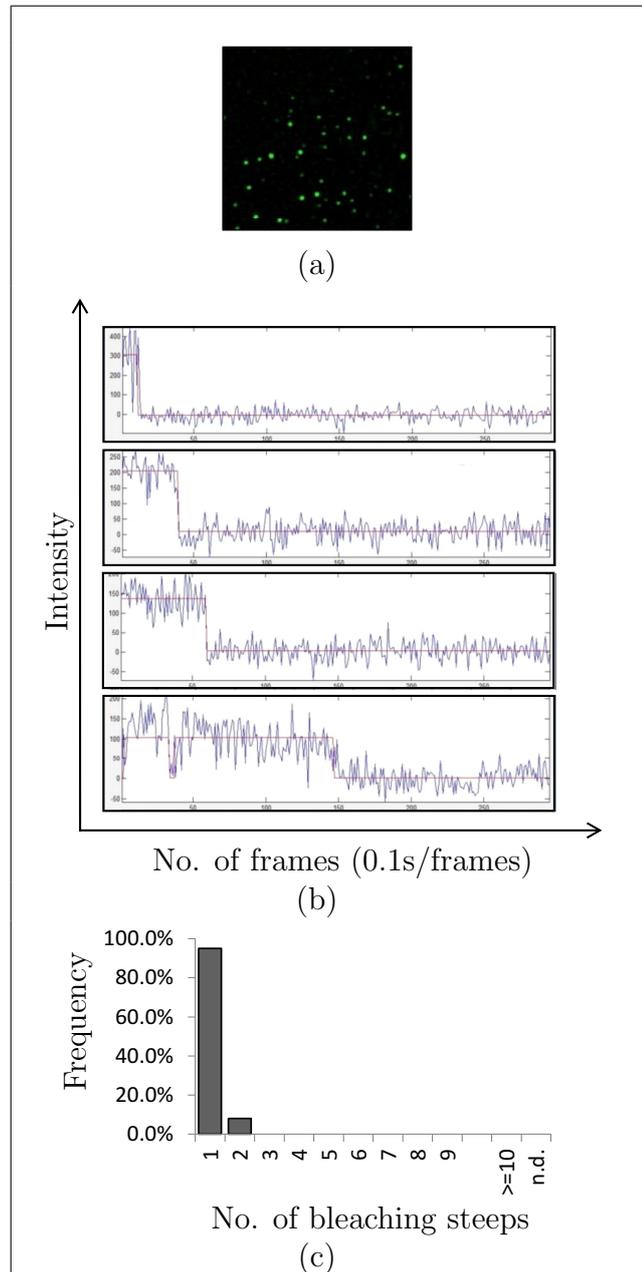


Figure 5.4: Testing the purity and the state of GFP only. **(a)** An image of GFP spots under the microscope. This image was analysed via MatLab, for more details see Chapter 3 and [Weinmeister, 2014]. **(b)** Time traces results for GFP. Traces is showing one bleaching step of GFP is exist. **(c)** Bar Chart of GFP results, the X axis shows the distribution state of GFP spots and the Y axis shows the percentage of GFP. No pre-mRNA was added to this experiments.

## 5.1 Introduction

---

Moreover, the proportion of spots in my experiments with GloC in the early complex E and complex A conditions was investigated in the matter if the proportion does contain two or more RNA molecules. Figure 5.4 shows almost single bleaching step appears when calculating the un-colocalised spots (Cy5-RNA).

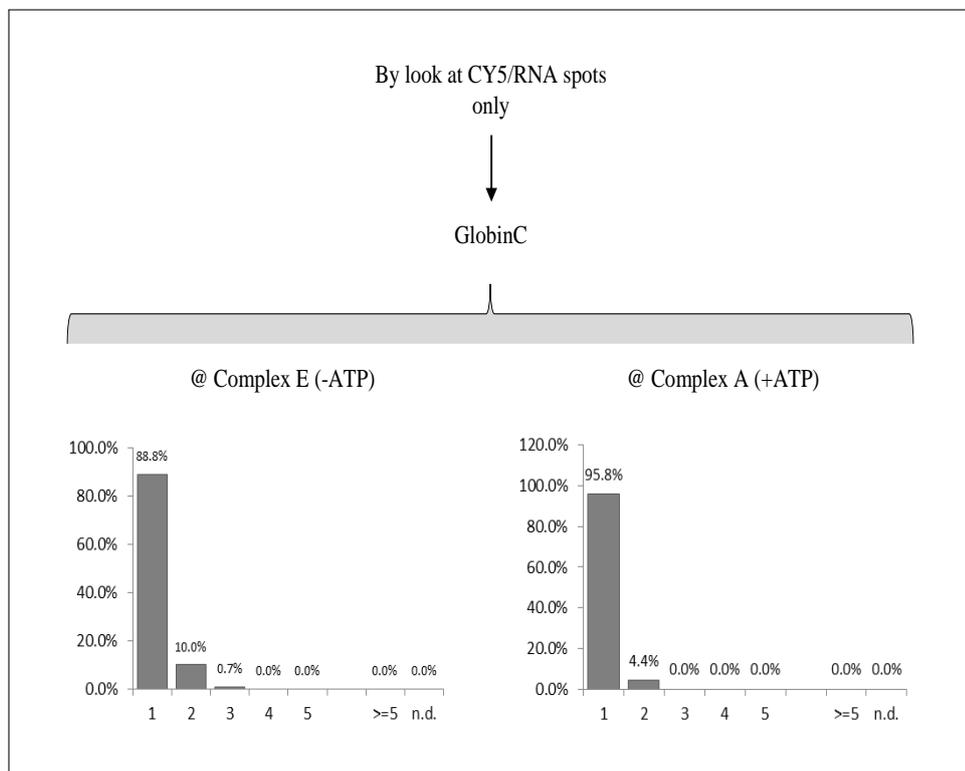


Figure 5.5: Testing the proportion of un-colocalised spots (Cy5-RNA). Left chart bar indicate a single molecule experiment in the absence of ATP. Right chart bar indicate a single molecule experiment in the presence of ATP and  $\alpha$ U6 oligo. The X axis shows the distribution state of the un-colocalised spots of Cy5-RNA. The Y axis shows the percentage of the un-colocalised spots of Cy5-RNA. The pre-mRNA used in this experiment is GloC.

As the binding distributions could be affected by the way that splicing is stalled at different complexes, multiple experiments were done and compared by Dr. Li Chen [Li, 2016] as mentioned previously in this work, see (Chapter

## 5.1 Introduction

3). All single molecule experiments were performed by actually stalling the complex formation through the absence of ATP, the addition of interfering oligos or mutations.

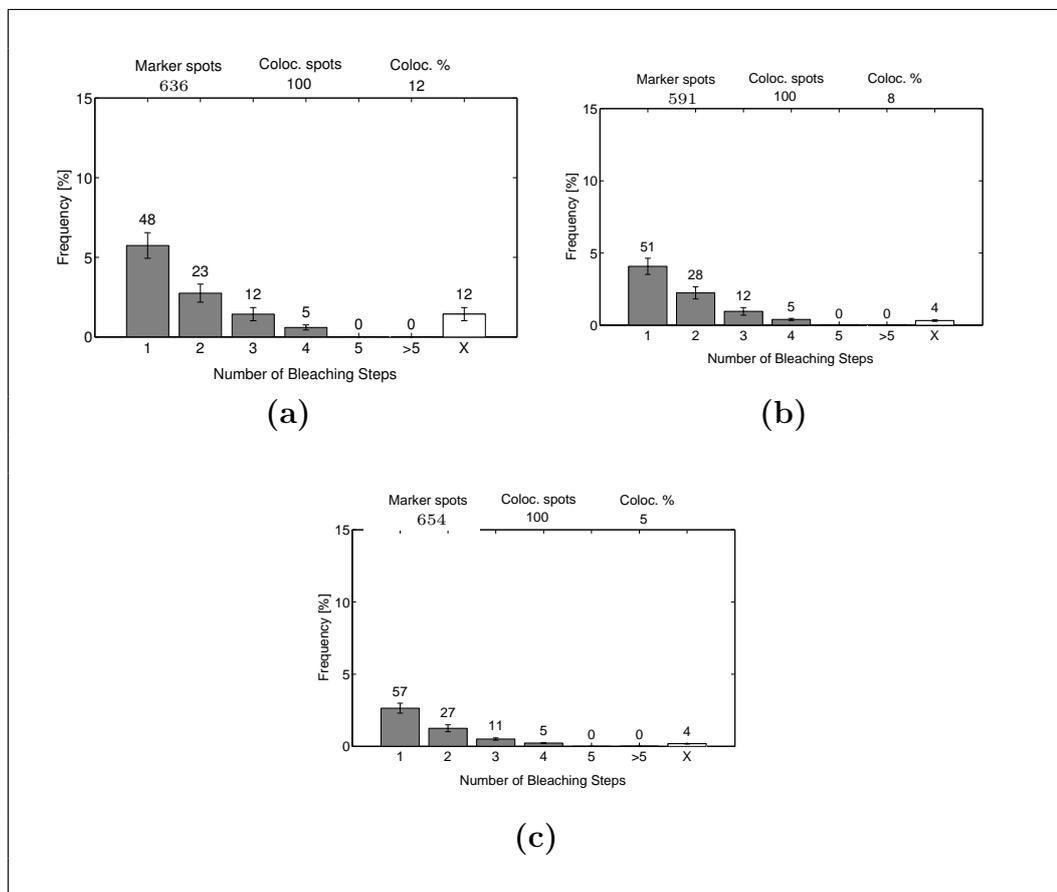


Figure 5.6: Bar chart of the amount of SRm160 present in the nuclear extract. **(a)** GloC pre-mRNA has incubated with SRm160 at the early complex E condition. **(b)** GloC pre-mRNA has incubated together with an abundance of commercial nuclear extract in a ratio 1:1 at the early complex E condition. **(c)** GloC pre-mRNA has incubated together with an abundance of commercial nuclear extract in a ratio 1:2 at the early complex E condition. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

## 5.2 SRm160 in Different Complexes; *the scope from different constructs*

---

Additionally, the amount of SRm160 present in the home made nuclear extract may play a role influencing the number of SRm160 binding. Experiments with an abundance of commercial nuclear extract to lower the ratio of SRm160 were done to compare with experiments with an abundance of SRm160 in a nuclear extract made from 293T cell for a higher ratio. The relative abundance of commercial nuclear extract leads to the same binding distribution whereas the abundance of SRm160 lead to a high colocalisation with binding dominated by large accumulations of SRm160, see figure 5.6.

Furthermore, the unspecific binding of SRm160 to the glass surface of the sample chamber might interfere with the observed results. The top of a glass chamber might have to cover with PEG or BSA. Experiments to cover the slide were performed and compared by Dr. Robert Weinmeister [Weinmeister, 2014]. The obtained results showed that the presence and the absence of PEG or BSA result similar binding distributions.

## 5.2 SRm160 in Different Complexes; *the scope from different constructs*

### 5.2.1 GloC Construct

The GloC construct was used to understand the state of SRm160 binding to a pre-mRNA and to assess the splicing reaction at single molecule level at the early complex E or at complex A conditions, see figure 5.7.

#### 5.2.1.1 GloC (Non-Modified Construct) at the Early Complex E (the depletion of ATP)

In the splicing reaction, the early complex E is formed in the absence of ATP and U1 snRNP that binds the RNA construct at the 5' splice site (5'

## 5.2 SRm160 in Different Complexes; *the scope from different constructs*

---

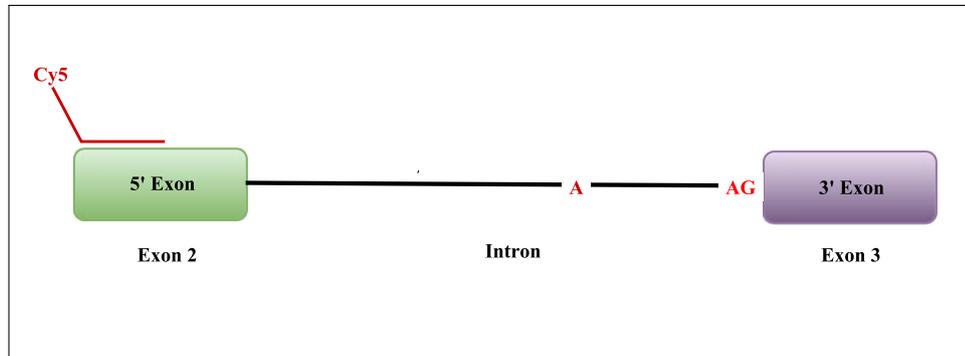


Figure 5.7: Schematic diagram of GloC pre-mRNA construct. From the left; green box represents 5' exon, solid black line represents the intron and the purple box represents the 3' exon. Solid red line indicates the Cy5-labelled oligonucleotide.

ss). By this stage, pre-mRNA commits to splice. Previous studies identified that the 3' ss is already recognized by U2AF at the early complex E ; and U2 snRNP also found to binds to pre-mRNA weakly through protein-protein interactions . To investigate the state of SRm160, how it bind pre-mRNA during those reactions and further to do single molecule analysis of complexes, single molecule experiment is implemented using a home made nuclear extract (SRm160). This extract was first tested in splicing reaction to ensure the validity and activity in *in vitro* splicing reaction, see (Chapter 4). Splicing was stalled at the early complexes E or at complex A to accumulate species of predominantly one complex. Stalling the splicing in the early complex E was achieved by depletion of ATP, while stalling at complex A was by the addition of ATP and a 2'-o-methyl oligonucleotide complementary to the 3' end of U6 snRNP (anti-U6 oligo). A pre-mRNA with a consensus 5' splice site (GloC) was used. The pre-mRNA construct was annealed to Cy5 oligonucleotide to ensure the visibility under the microscope, see (Chapter 4). After incubation in extract pre-incubated to deplete ATP, 12% of the RNA complexes contained

## 5.2 SRm160 in Different Complexes; *the scope from different constructs*

---

multiple molecules of GFP-SRm160 were identified. Figure 5.8 shows the frequencies of GloC pre-mRNA molecules vs. bleaching steps of colocalized mGFP-SRm160 in one, two, three or by a multitude of them visible as a long tail. These multiple binding could be referred to as aggregates.

### 5.2.1.2 Globin C Construct at Complex A (the addition of ATP)

The splicing complex formed after the formation of the early complex E is complex A. This complex is formed by base pairing of U2 snRNA with the branch point (BP) of pre-mRNA in the presence of ATP. To investigate the state of SRm160 binding in complex A, the splicing reaction was stalled at complex A condition first. The stalling reaction was via the addition of ATP and an oligonucleotide complementary to U6 snRNP. The result pattern is dramatically different. The binding of SRm160 is increased significantly compared to the early complex E condition (Figure 5.8). For example, in GloC, the percentages of SRm160 bound at the early complex E condition is 12%.

However, the percentage is increased up to 21% at complex A. This indicates that SRm160 molecules might interact very weakly with the pre-mRNA. Strikingly, almost all complexes contain one to two molecules of GFP-SRm160. The level of colocalization was much higher than in the absence of ATP.

## 5.2 SRm160 in Different Complexes; *the scope from different constructs*

---

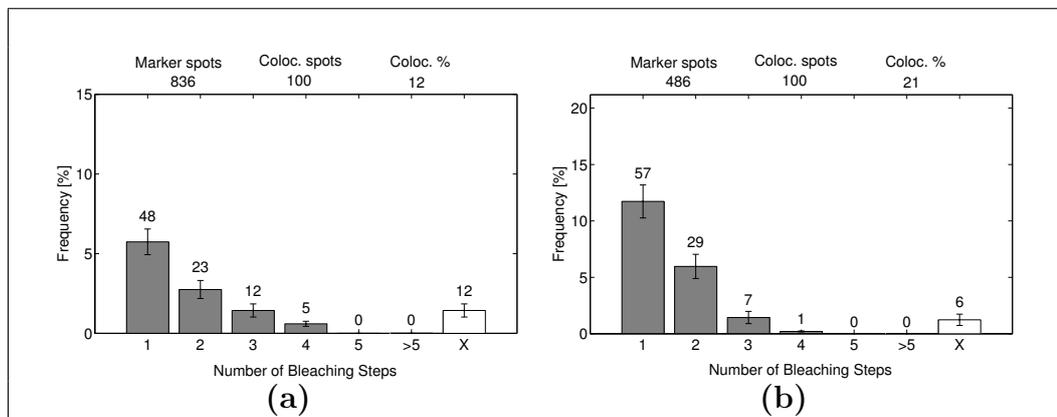


Figure 5.8: Bar chart of SRm160 binding to GloC pre-mRNA construct. Single molecule experiment shows the binding of SRm160 at complex A (in the presence of both ATP and anti-U6 oligo) in compare to the early complex E (no ATP). **(a)** At the early complex E. **(b)** At complex A (with ATP). Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

## 5.2 SRm160 in Different Complexes; *the scope from different constructs*

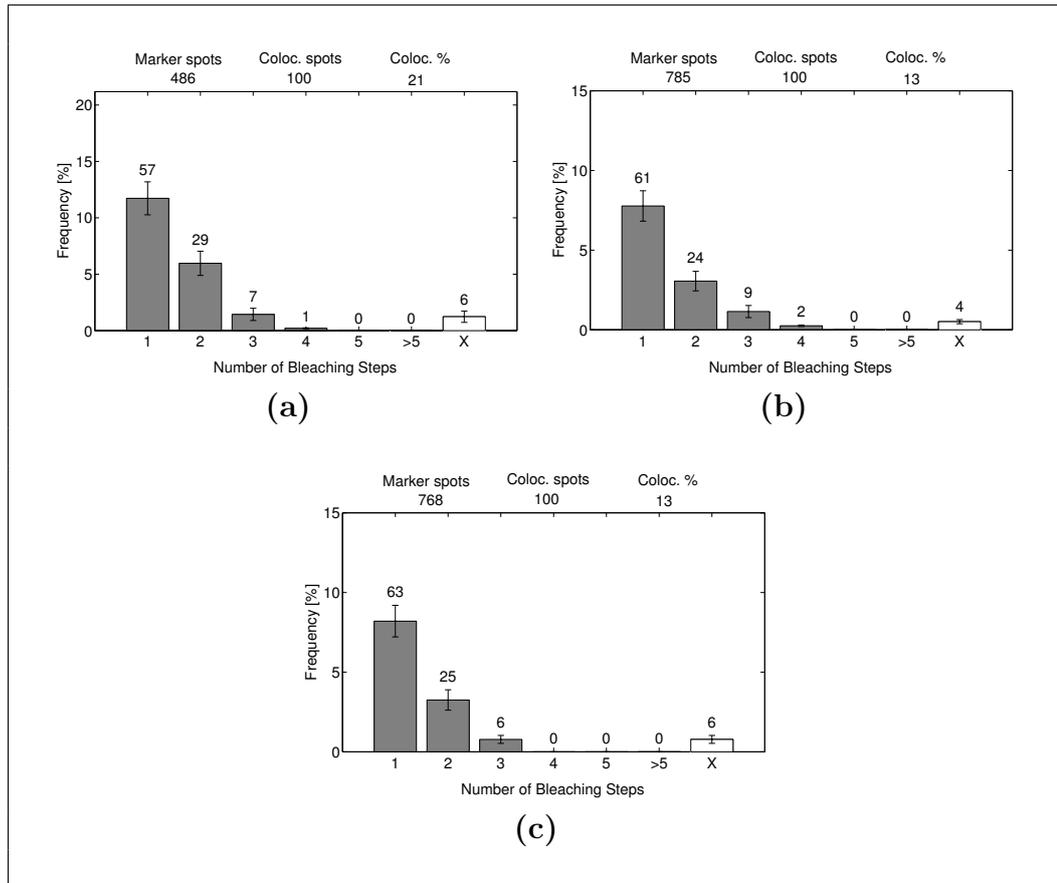


Figure 5.9: Bar chart of SRm160 binding to GloC pre-mRNA construct. Single molecule experiment shows the binding of SRm160 at complex A condition (with ATP) in the presence of different anti oligos. **(a)** At complex A (with ATP). **(b)** At complex A in the presence of anti-U1 oligo. **(c)** At complex A in the presence of anti-U2 oligo. Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

## 5.2 SRm160 in Different Complexes; *the scope from different constructs*

---

### 5.2.2 The Interplay of SRm160 Protein With the U1 snRNP and the 5' Splice Site; Globin M

For deep investigation if whether the SRm160 molecules in the presence and the absence of ATP is the result of interactions with U1 snRNP; Globin M (the consensus 5' splice site of the GloC pre-mRNA was inactivated by mutation into a non-splice site sequence), see figure 5.10. By using this substrate, no binding of U1 snRNP should occur without a 5' splice site present. Gel electrophoresis shows that Globin M substrate does not splice because of the absence of the consensus 5' splice site.

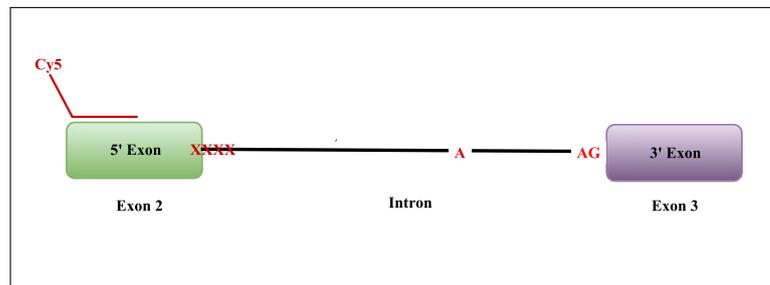


Figure 5.10: Schematic diagram of Glo M construct From the left; green box represents 5' exon, solid black line represents the intron and the purple box represents the 3' exon. Solid red line indicates the Cy5-labelled oligonucleotide. Multiple red label (X) represents the absence of the corresponding site.

The binding distribution of SRm160 molecules in condition supported the formation of the early complex E has no clear definition and is dominated by accumulations of SRm160 proteins at the pre-mRNA (Figure 5.11). In the presence of ATP and an oligonucleotide complementary to U6 snRNA ( $\alpha$ U6), a significant increase towards one molecule of SRm160 bound to RNA (Figure 5.11). Although Globin M does not form complex A due to the lack of a 5' splice site, incubation with ATP does lead to multiple remaining molecules of SRm160 bound to pre-mRNA. Regarding the distribution pattern, very similar

## 5.2 SRm160 in Different Complexes; *the scope from different constructs*

results occurred in comparison to GloC, E3 with high colocalisation efficiency.

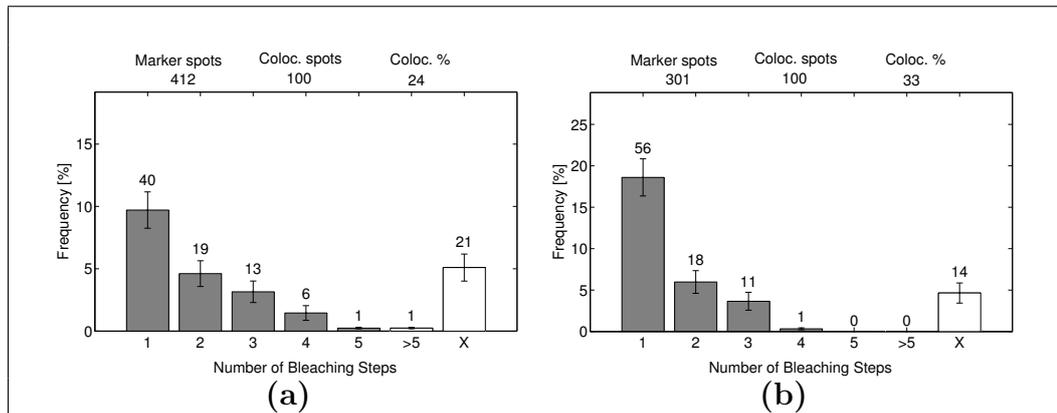


Figure 5.11: Bar chart of SRm160 binding with Globin M pre-mRNA under different conditions. Single molecule experiment shows the binding of SRm160 at the early complex E (no ATP) and at complex A (the addition of both ATP and anti-U6 oligo). **(a)** At the early complex E. **(b)** At complex A. Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

### 5.2.3 Intron-less Globin C; mRNA

As from all previous studies, the association of SRm160 to GloC pre-mRNA depends on the presence of U2 snRNP. If this is the case, presence of splicing signals could affect the interaction of GFP-SRm160 which indicates that the protein of SRm160 has no binding to the intron-less construct. To draw a complete splicing cycle, it is interesting to further investigate whether SRm160 will follow the logic of this binding and still bind to the intron-less substrate (spliced mRNA), see figure 5.12. To determine this, spliced GloC was used. Interestingly, single molecule colocalisation experiments showed unexpected

## 5.2 SRm160 in Different Complexes; *the scope from different constructs*

---

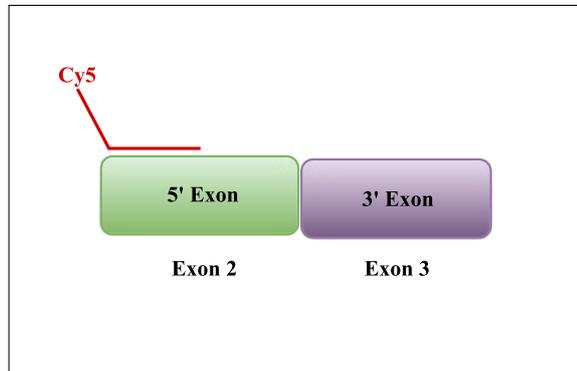


Figure 5.12: Schematic diagram of spliced GloC RNA. From the left; green box represents 5' exon and the purple box represents the 3' exon. Solid red line indicates the Cy5-labeled oligonucleotide.

result in which SRm160 proteins are still bind to the intron-less GloC construct in the early complex E and complex A condition, see Figure 5.13.

To follow this, intron-less GloC mRNA was made by mutagenesis PCR of the GloC plasmid, followed by PCR amplification and in vitro transcription. At the beginning we were expecting that no SRm160 would be bound to GloC intron-less RNA as there are no intron sequences. However, unexpectedly, SRm160, still bound to intron-less GloC in the absence and the presence of ATP (Figure 5.13). In the absence of ATP, the distribution pattern of SRm160 bound the spliced RNA was very similar to GloC with high percentage of colocalization efficiency. Addition of ATP and an oligonucleotide complementary to U6 oligo indicates similar results to GloC with slight decrease in the level of the second GFP-SRm160. These results support the idea that the presence of U2 snRNP is important to recruit GFP-SRm160 molecules. Additionally, The presence of this binding of SRm160 molecules might be due to the presence of ESEs sites in the exons which consists with the results in section that the binding of SRm160 is influenced by the presence of ESEs sites.

### 5.3 SRm160 and the Splicing Enhancer; SMN2

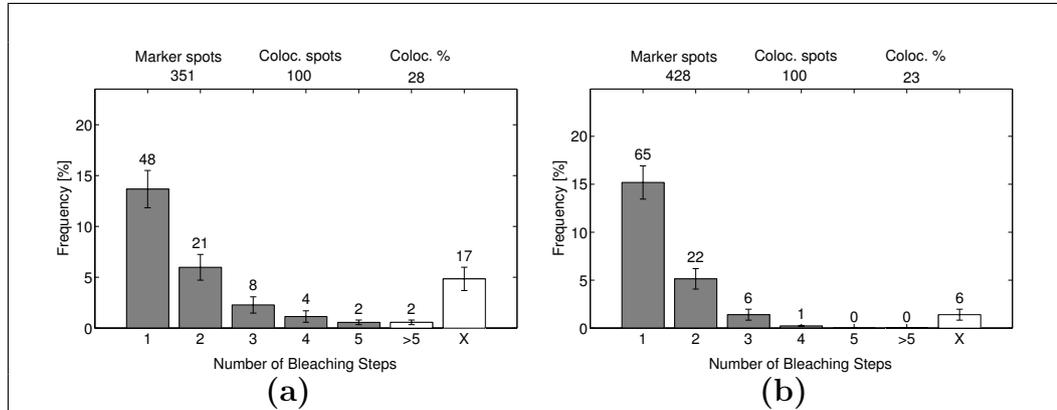


Figure 5.13: Bar chart of SRm160 binding with spliced mRNA (intronless) under different conditions. Single molecule experiment shows the binding of SRm160 at the early complex E (no ATP) and at complex A (with ATP) in the presence of anti-U6 oligo. **(a)** At the early complex E. **(b)** At complex A. Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

### 5.3 SRm160 and the Splicing Enhancer; SMN2

The ESE sequences are important for the recognition of splice site in pre-mRNA. These sequences are bound by serine-arginine (SR) repeat proteins that promote the assembly of splicing complexes at specific splice sites.

Regarding SRm160 protein, the number of binding protein of SRm160 is generally seen to be increased as the splicing enhancer sites are increased. This relation was demonstrated to be influenced by a specific sequences, such as GAA-repeats sites. In this research work, a construct named as SMN2 with and without ESE sequences were used in order to investigate if an effect of the ESE sites can be seen on the observed binding pattern of SRm160 proteins or

### 5.3 SRm160 and the Splicing Enhancer; SMN2

---

not. This substrate is based on GloC construct which derived from exon 2 of GloC and exon 7 from SMN2 (Figure 5.14).

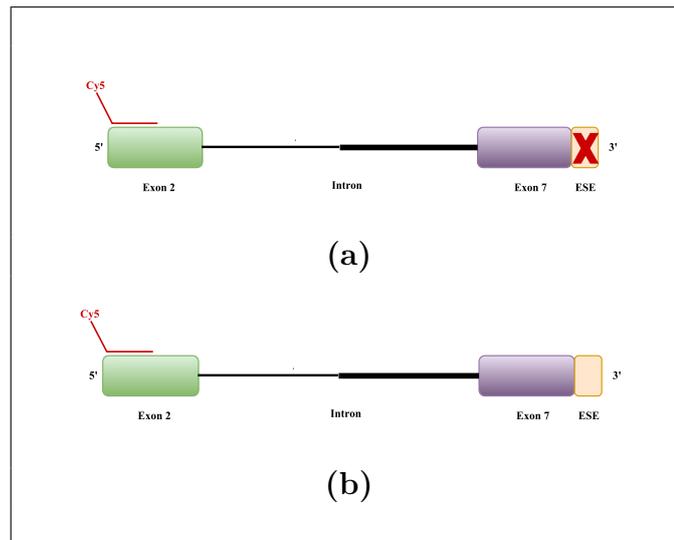


Figure 5.14: Schematic diagram of SMN2 substrate. **(a)** Represents SMN2 construct without the ESE sequence. **(b)** Represents SMN2 construct without the ESE sequence. From the left; green box represents 5' exon, solid black line represents the intron and the purple box represents the 3' exon. Solid red line indicates the Cy5-labeled oligonucleotide. Red label (X) represents the absence of the site.

In the absence of ATP, under the conditions allowing the early complex E assembly, pre-mRNA associates with multiple number of GFP-SRm160 molecules. The number of bound GFP-SRm160 is strikingly increased with increasing in colocalisation percentage (from 13% to 24%) in SMN2 containing ESE sequence (Figure 5.15). In the presence of ATP and an oligonucleotide complementary to U6 snRNP (anti-U6 oligo), binding of GFP-SRm160 is restricted to two defined molecules in the presence of ESE and 1 to 2 molecules in the absence of ESE. These results indicate that the number of GFP-SRm160 that bind pre-mRNA substrate is highly depend on the ESE sequence (Figure 5.15).

### 5.3 SRm160 and the Splicing Enhancer; SMN2

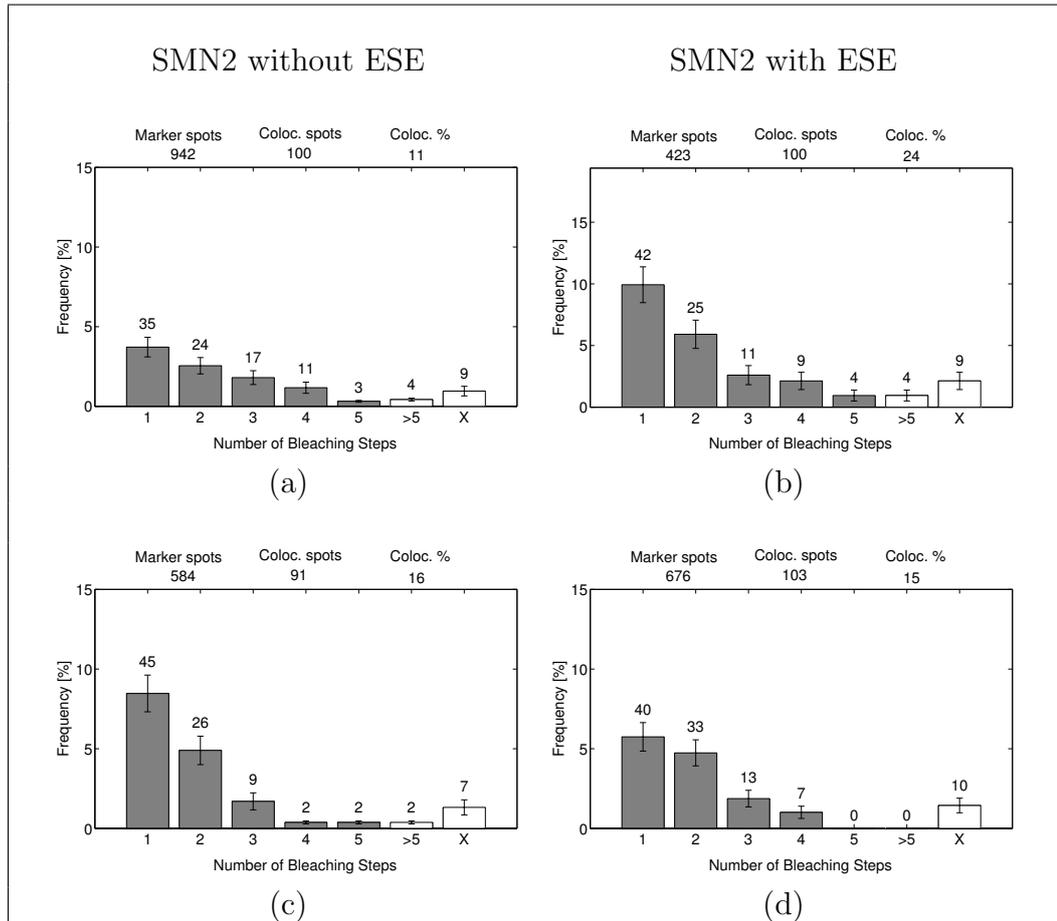


Figure 5.15: Bar chart of SRm160 association with SMN2 pre-mRNA under different conditions. Left column represents SMN2 without ESE sequence. Right column represents SMN2 with the ESE sequence. **(a)** and **(b)** Represent the frequencies of SRm160 binding to SMN2, showing the binding of mGFP-SRm160 in condition supporting the formation of the early complex E (no ATP). The level of GFP-SRm160 bound to pre-mRNA is higher in the presence of ESE. **(c)** and **(d)** Binding of SRm160 molecules to SMN2 at complex A. Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

### 5.3 SRm160 and the Splicing Enhancer; SMN2

---

Moreover, the addition of phoSTOP to a condition supported the formation of the early complex E showed (Figure 5.16) does not show a significant changes regarding the binding of GFP-SRm160. The obtained results indicates that the phoSTOP decrease the recruiting of GFP-SRm160 to SMN2 substrates in both the absence and the presence of ESE sites.

In addition, to determine whether U1 and U2 snRNPs contribute to the binding of SRm160, SMN2 with and without ESE sequences were used in extract pre-incubated reaction to deplete ATP in the presence of an oligonucleotide complementary to U1 snRNP. Results demonstrated that the association of GFP-SRm160 with SMN2 substrate does not depend on U2 snRNP even in the presence of an ESE sequence (Figure 5.17). Similar distribution pattern with low colocalisation efficiency were observed in the absence of this oligonucleotide (Figure 5.17). These results are arguing the previous observation that the SRm160 does not bind to pre-mRNA in the absence of U2 snRNP. Surprisingly, a single GFP-SRm160 molecule was identified when an oligonucleotide complementary to U2 snRNP existed. These results demonstrate various points; the association of SRm160 with SMN2 pre-mRNA is enhanced via ESE sequences, highly depend on U2 snRNP but not U1 snRNP.

### 5.3 SRm160 and the Splicing Enhancer; SMN2

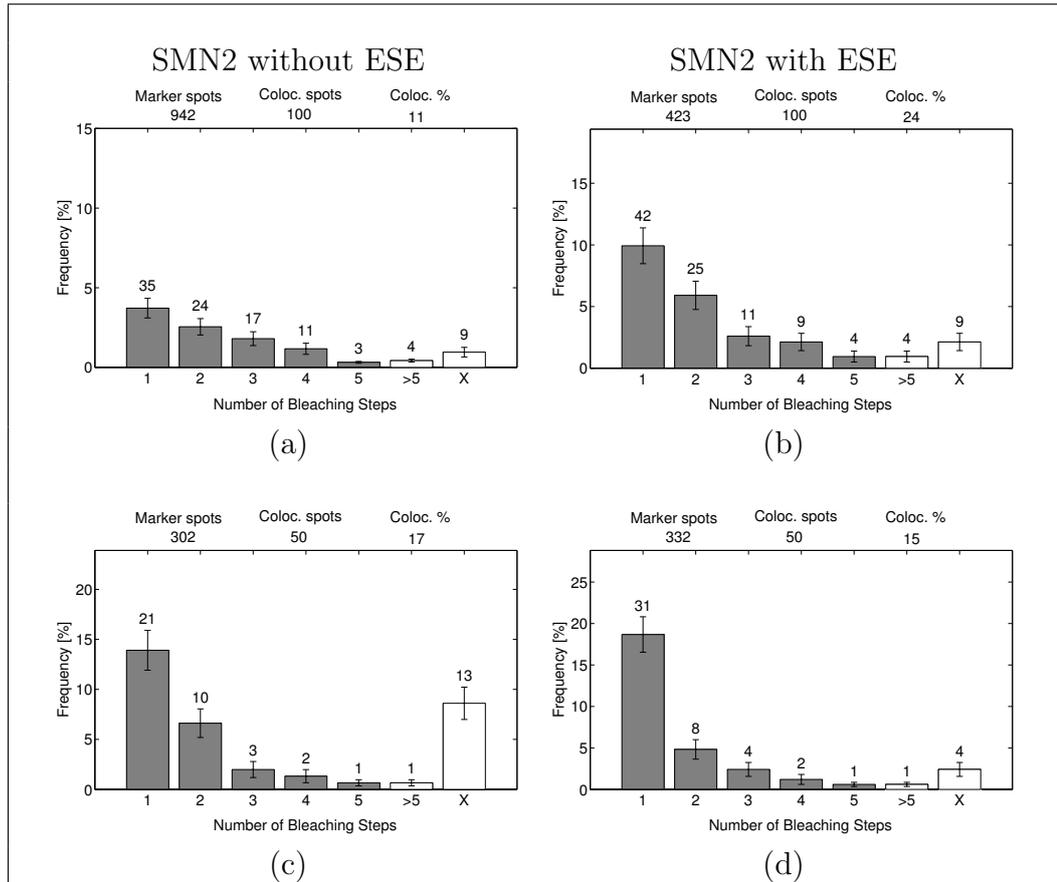


Figure 5.16: Bar chart of SRm160 association with SMN2 pre-mRNA under different conditions. Left column represents SMN2 without ESE sequence. Right column represents SMN2 with the ESE sequence. **(a)** and **(b)** Frequencies of SRm160 binding to SMN2, showing the binding of mGFP-SRm160 in condition supporting the formation of the early complex E (no ATP). **(c)** and **(d)** Using PhosStop in the absence of ATP. Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

### 5.3 SRm160 and the Splicing Enhancer; SMN2

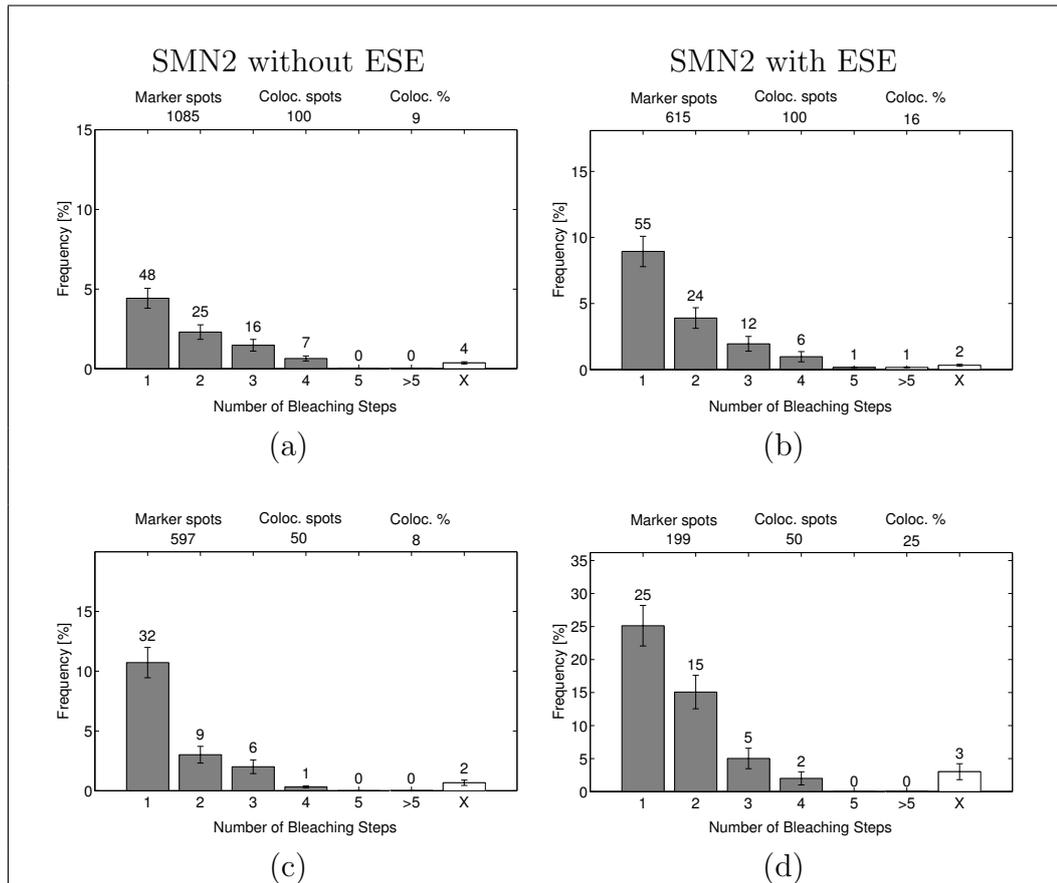


Figure 5.17: Bar chart of SRm160 association with SMN2 pre-mRNA under different conditions. Left column represents SMN2 without ESE sequence. Right column represents SMN2 with the ESE sequence. **(a)** and **(b)** At the early complex E (no ATP) in the presence of anti-U1 oligo showing a slight shift in the number of SRm160 towards one binding step with low colocalized efficiency. **(c)** and **(d)** At the early complex E (no ATP) in the presence of both anti-U1 oligo and phoSTOP. Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

### 5.3 SRm160 and the Splicing Enhancer; SMN2

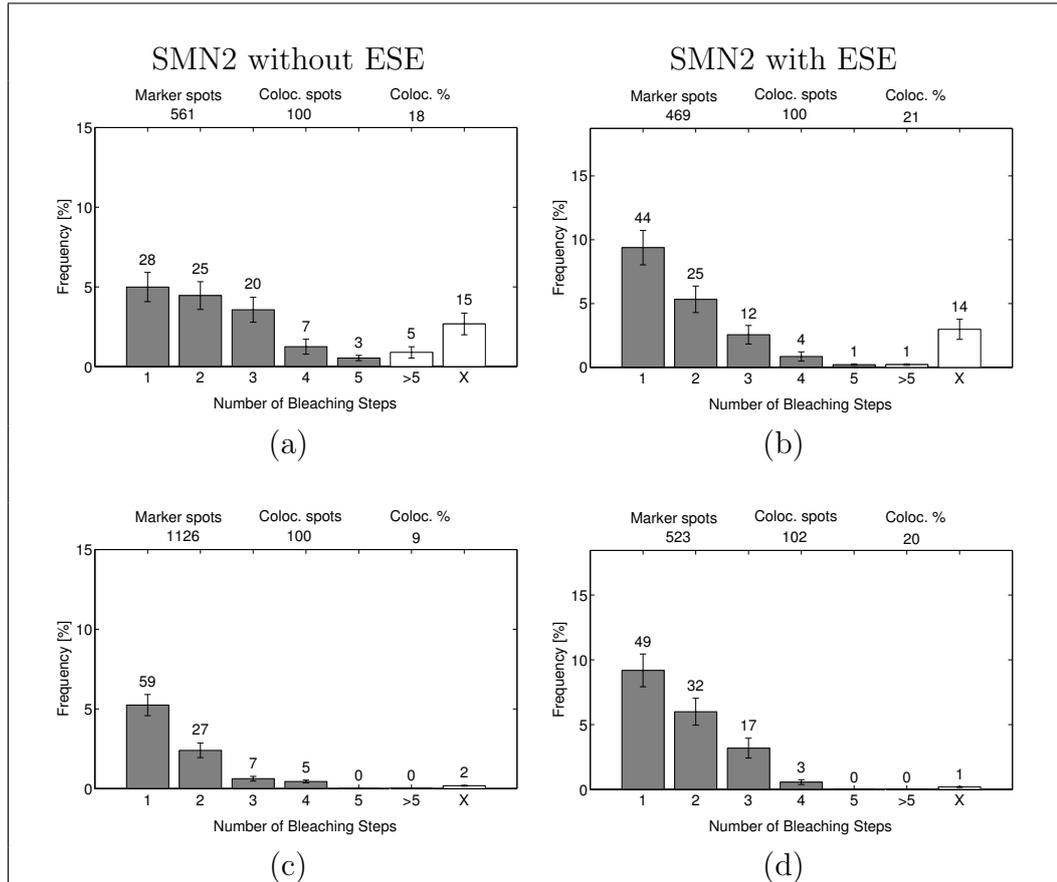


Figure 5.18: Bar chart of SRm160 association with SMN2 pre-mRNA under different conditions. Left column represents SMN2 without ESE sequence. Right column represents SMN2 with the ESE sequence. (a) and (b) At complex A (with ATP) in the presence of anti-U1 oligo. (c) and (d) At complex A (with ATP) in the presence of anti-U2 oligo. Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The axis X shows the percentage of SRm160 proteins binding. The (X) axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

### 5.3 SRm160 and the Splicing Enhancer; SMN2

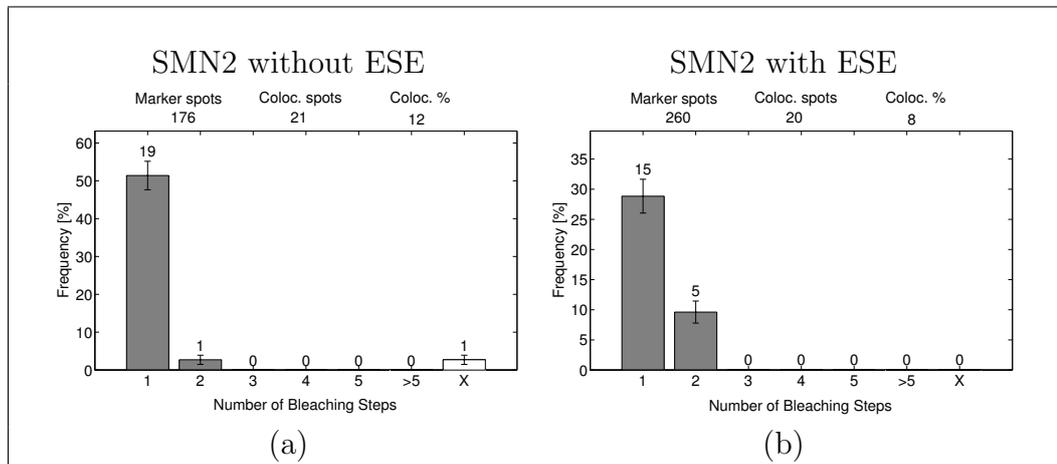


Figure 5.19: Bar chart of SRm160 association with SMN2 pre-mRNA under different conditions. Left column represents SMN2 without ESE sequence. Right column represents SMN2 with the ESE sequence. (a) and (b) At complex A (with ATP) in the presence of both anti-U1 and anti-U2 oligos. Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

## 5.4 SRm160 and the Tra2 Site

---

### 5.4 SRm160 and the Tra2 Site

In the previous studies, it was demonstrated that that hTra2 $\alpha$  and hTra2 $\beta$  of *Drosophila* Tra2 bind to ESEs containing GAA repeats and, in conjugation with SR family proteins such as SRm160, promote ESE-dependent splicing. If this is the case, does the absence of hTra2 $\beta$  in SMN2 pre-mRNA substrate affect the recruitment of GFP-SRm160 molecules. To determine this case, SMN2 substrate without both hTra2 $\beta$  and ESE sequences is used in comparison to SMN2 without hTra2 $\beta$  site containing ESE sequence. SMN2 pre-mRNA was annealed to CY5 oligonucleotide described previously in Chapter 4, (Figure 5.20).

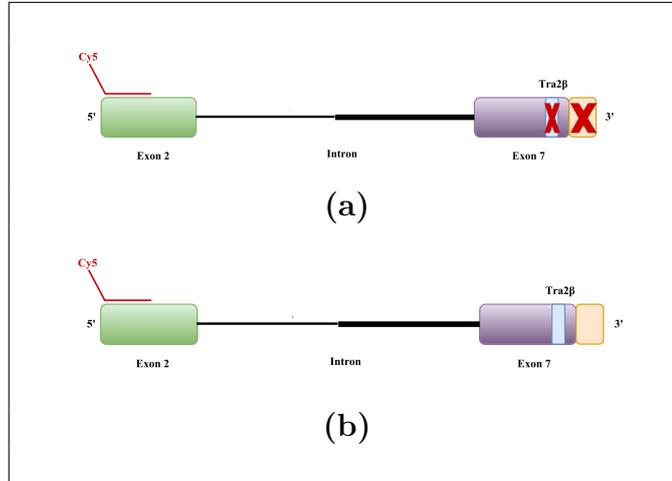


Figure 5.20: Schematic diagram of SMN2 substrate. (a) Represents SMN2 construct without both the ESE sequence and the hTra2 $\beta$  site. (b) Represents SMN2 construct with ESE sequence but without the hTra2 $\beta$  site. From the left; green box represents 5' exon, solid black line represents the intron and the purple box represents the 3' exon. Solid red line indicates the Cy5-labeled oligonucleotide. Red label (X) represents the absence of the corresponding site.

After incubation in nuclear extract pre-incubated to deplete ATP, 11% and 14% of the RNA complexes contained multiple molecules of GFP-SRm160

#### 5.4 SRm160 and the Tra2 Site

---

were identified in the absence and the presence of ESE sequence, respectively. Phostop treatment in the absence of ATP resulted in single defined GFP-SRm160 molecule and single molecule in the absence of the ESE sequence, and one to two molecules in the presence of ESE sequence (Figure 5.21). These results indicate that the multiple binding could be an aggregates or nonspecific binding of GFP-SRm160. Figure shows the frequencies of SMN2 pre-mRNA molecules vs. bleaching steps of colocalized mGFP-SRm160 in 1, 2, 3 etc. steps. In the presence of ATP, and an oligonucleotide complementary to U6 snRNP that stalls assembly at complex A, Hodson showed that the presence of ESE sequence recruited more GFP-SRm160 molecules Hodson et al. [2012]. Two GFP-SRm160 molecules were identified. The colocalization efficiency was increased from 9% in the absence of ESE to 19% in the presence of the ESE sequence (Figure 5.22). These results indicates that the presence of ESE sequences recruited more GFP-SRm160 in SMN2 substrate. This arguing with the previous studies that the association of SRm160 protein is specific for ESE-dependent splicing containing GAA repeats.

Strikingly, the addition of a nucleotide which is complementary to U1 snRNP to ATP-depleted reaction, the binding of SRm160 to SMN containing ESE sequences is slightly higher with similar distribution pattern to SMN2 in the absence of this oligonucleotide (Figure 5.23). These results indicate the absence of U1 snRNP does not affect the binding of GFP-SRm160 to pre-mRNA. Whereas in the absence of ESE sequence, one to two GFP-SRm160 molecules were identified. This supports the idea that level of GFP-SRm160 is enhanced by the presence of ESE sequences even in the presence of an oligonucleotide complementary to U1 snRNA. Surprisingly, the presence of an oligonucleotide complementary to U2 snRNA significantly reduced the number

#### 5.4 SRm160 and the Tra2 Site

---

of SRm160 molecule to a single molecule in the absence of ATP. This result in conjugation with the low colocalization efficiency suggest that the binding of SRm160 is associated with U2 snRNP.

Briefly, in this research, it was demonstrated that the number of GFP-SRm160 is enhanced via ESE sequences. The presence of an oligonucleotide complementary to U1 snRNA identified no effect on the number of GFP-SRm160 bound to pre-mRNA. Whereas the presence of an oligonucleotide complementary to U2 represents a significant reduction in the number of GFP-SRm160 molecules bound pre-mRNA. Integrally, the association of GFP-SRm160 to pre-mRNA does not stimulated by Tra2 site.

## 5.4 SRm160 and the Tra2 Site

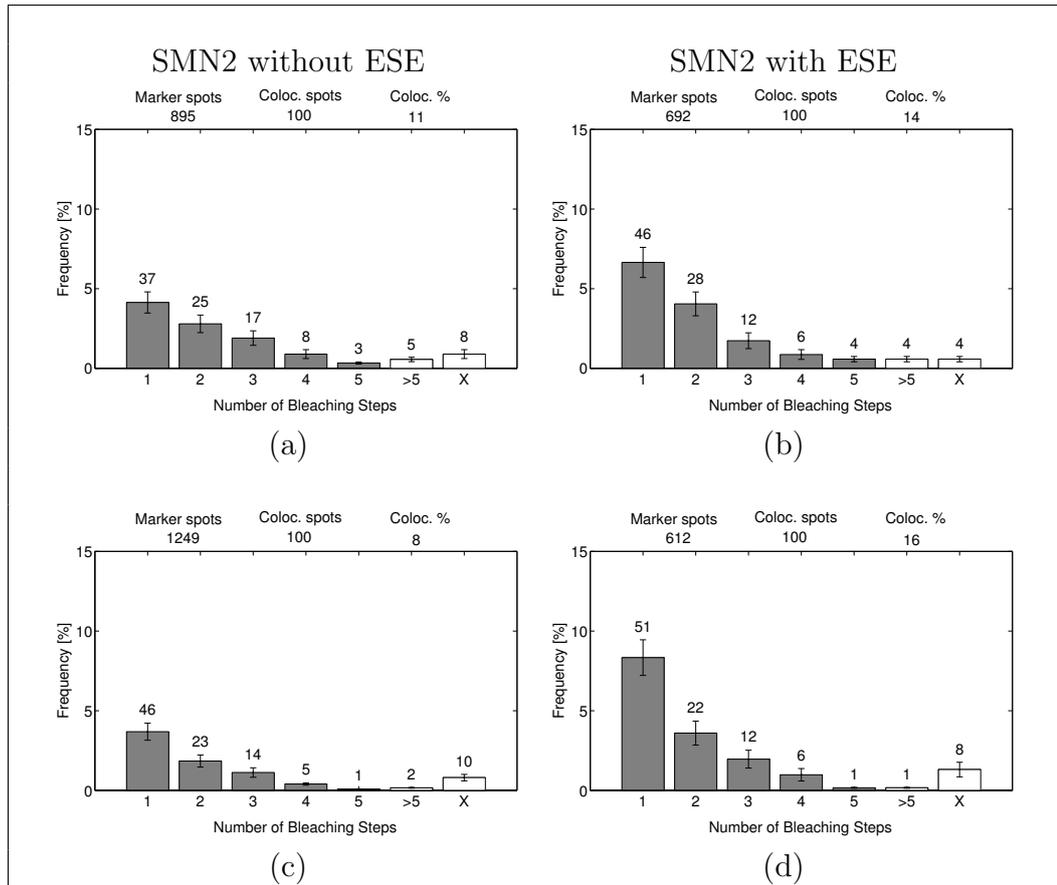


Figure 5.21: Bar chart of SRm160 association with SMN2 under different conditions. Left column represents SMN2 without both Tra2 and ESE sequences. Right column represents SMN2 without Tra2 but with the ESE sequence. (a) and (b) Frequencies of SMN2 showing the binding of mGFP-SRm160 in condition supporting the formation of the early complex E (no ATP). (c) and (d) Using PhosStop in the absence of ATP. Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

## 5.4 SRm160 and the Tra2 Site

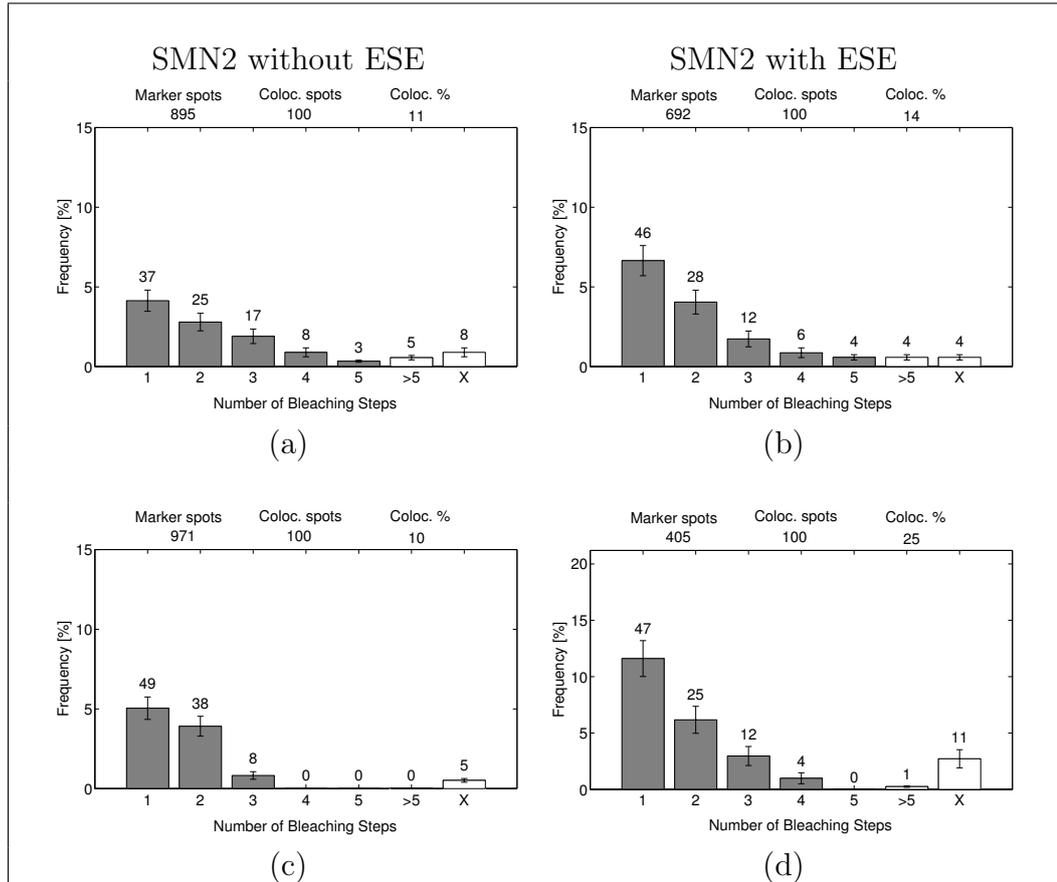


Figure 5.22: Bar chart of SRm160 association with SMN2 under different conditions. Left column represents SMN2 without both Tra2 and ESE sequences. Right column represents SMN2 without Tra2 but with the ESE sequence. (a) and (b) Frequencies of SMN2 showing the binding of mGFP-SRm160 in condition supporting the formation of the early complex E (no ATP). (c) and (d) At complex A (with ATP). Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

## 5.4 SRm160 and the Tra2 Site

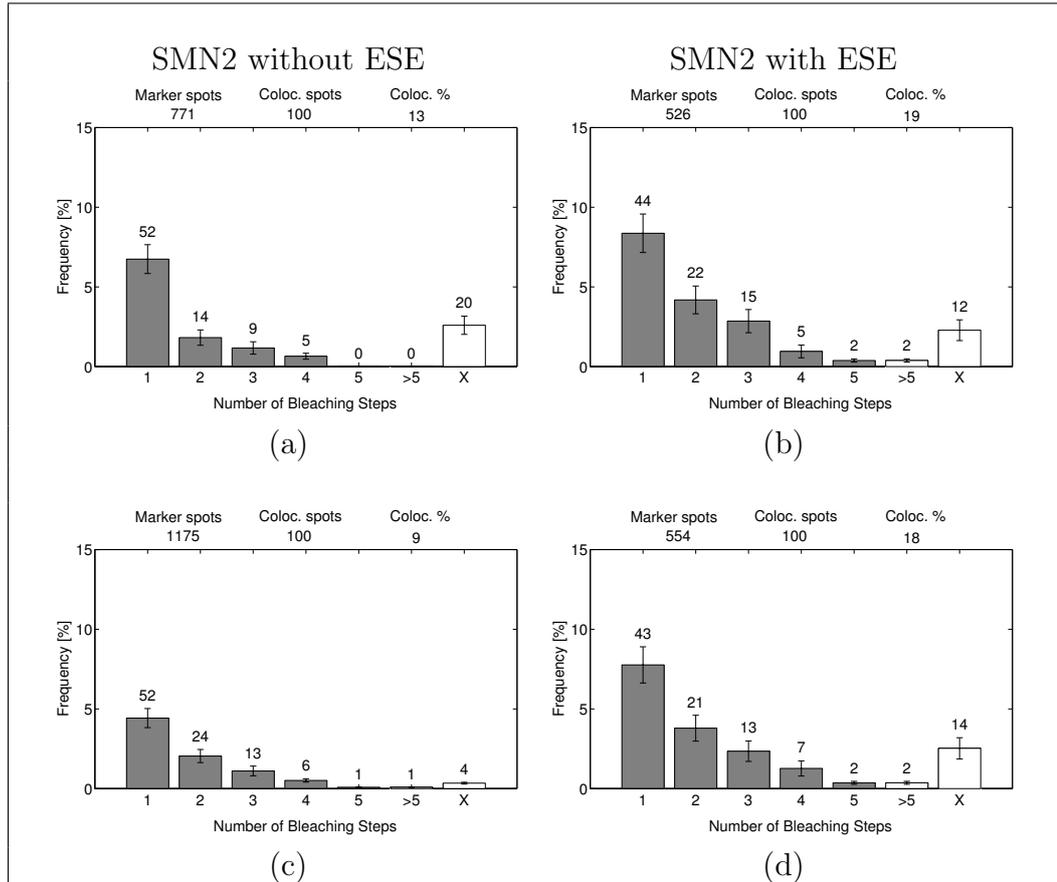


Figure 5.23: Bar chart of SRm160 association with SMN2 under different conditions. Left column represents SMN2 without both Tra2 and ESE sequence. Right column represents SMN2 without Tra2 but with the ESE sequence. (a) and (b) At the early complex E (no ATP) in the presence of anti-U1 oligo. (c) and (d) At the early complex E (no ATP) in the presence of both phoSTOP and anti-U1 oligo. Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation.

## 5.4 SRm160 and the Tra2 Site

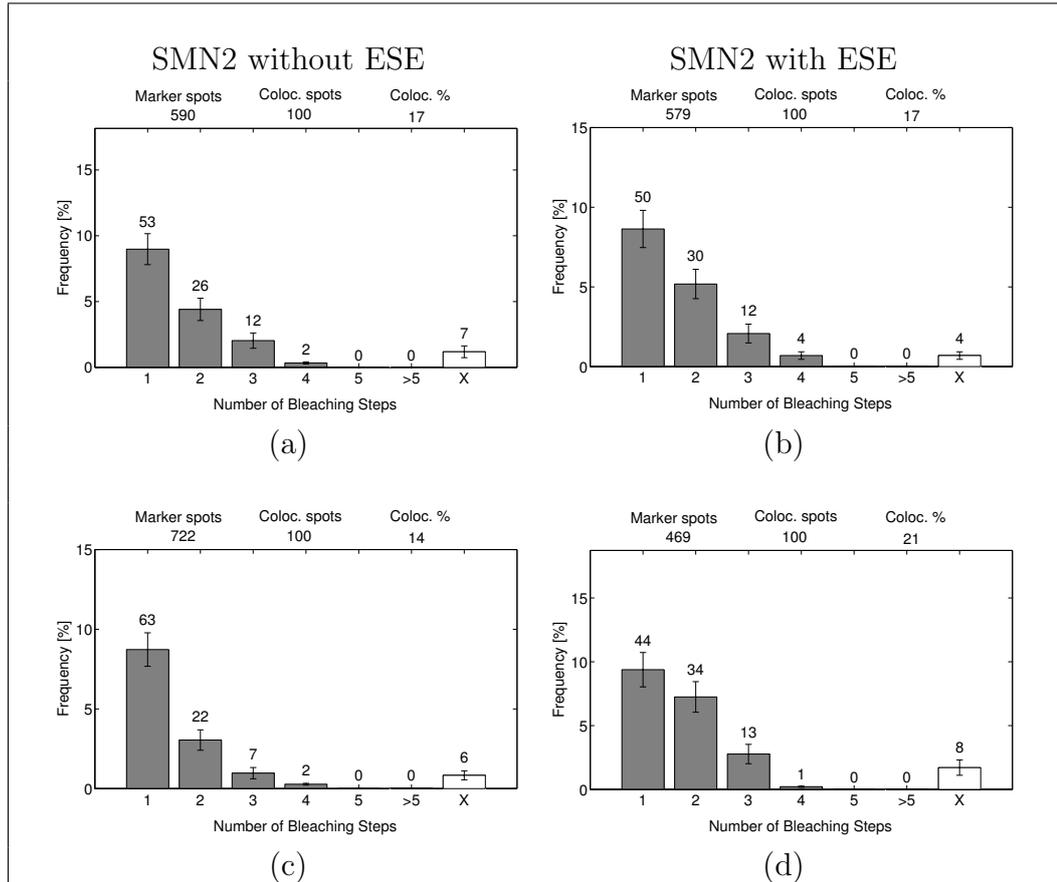


Figure 5.24: Bar chart of SRm160 association with SMN2 under different conditions. Left column represents SMN2 without both Tra2 and ESE sequence. Right column represents SMN2 without Tra2 but with the ESE sequence. **(a)** and **(b)** At complex A (with ATP) in the presence of anti-U1 oligo. **(c)** and **(d)** At complex A (with ATP) in the presence of anti-U2 oligo. Single GFP-SRm160 was identified. The number of bleaching steps indicates the number of mGFP-SRm160 molecules bound to each molecule of RNA. The X axis shows the distribution of this binding. The Y axis shows the percentage of SRm160 proteins binding. On the top of the chart; marker spots identify the number of marker molecules (Cy5-labelled pre-mRNA), Coloc. spots identify the total number of SRm160 molecules and Coloc. % identify the percentage of the colocalisation. .

### 5.5 SRm160 Vs. the 5' exon

As all constructs used in this work consist of a strong 5' exon and one intron, it will be more interesting to investigate whether the appearance of a binding behaviour of SRm160 protein is depended on the presence of those ESEs in the 5' exon.

To establish the investigation of this hypothesis, a series of GloC substrates were used. Those substrates differed in the length of the 5' exons. These substrates were basically designed from GloC. The procedure of the preparation is as follows:

- a. Designing a proper primers that complement to the right length of the 5' exon site and to ensure annealing temperature to be around 55-60°C.
- b. PCR as in Chapter 2, section 2.1.1.1.
- c. Agarose gel electrophoresis as in Chapter 2, section 2.1.1.1.
- d. Purification as in Chapter 2, section 2.1.1.2.
- e. Checking the concentration using nanodrop.

After ensuring that all resulted fragments (designed oligonucleotides) are corresponding to the write size, standard splicing reaction were implemented (Chapter 2, section 2.1.3.1 to 2.1.3.2). The obtained results indicates that the ESEs are not important for the splicing of GloC. (Figure 5.25, 5.27, 5.28 and 5.29) as all substrates shows a splicing activity with different efficiencies (Figure 5.30, 5.31 and 5.32).

Overall, short exons substrates used here in this work shows a splicing signal whether ESEs exist or not. This findings raised a thought that the ESE

## 5.5 SRm160 Vs. the 5' exon

sites at the 5' exon might be a reason for bound-SRm160 molecules remaining on a substrate that lacks a 3' ESE or the intron or the 5' ss.

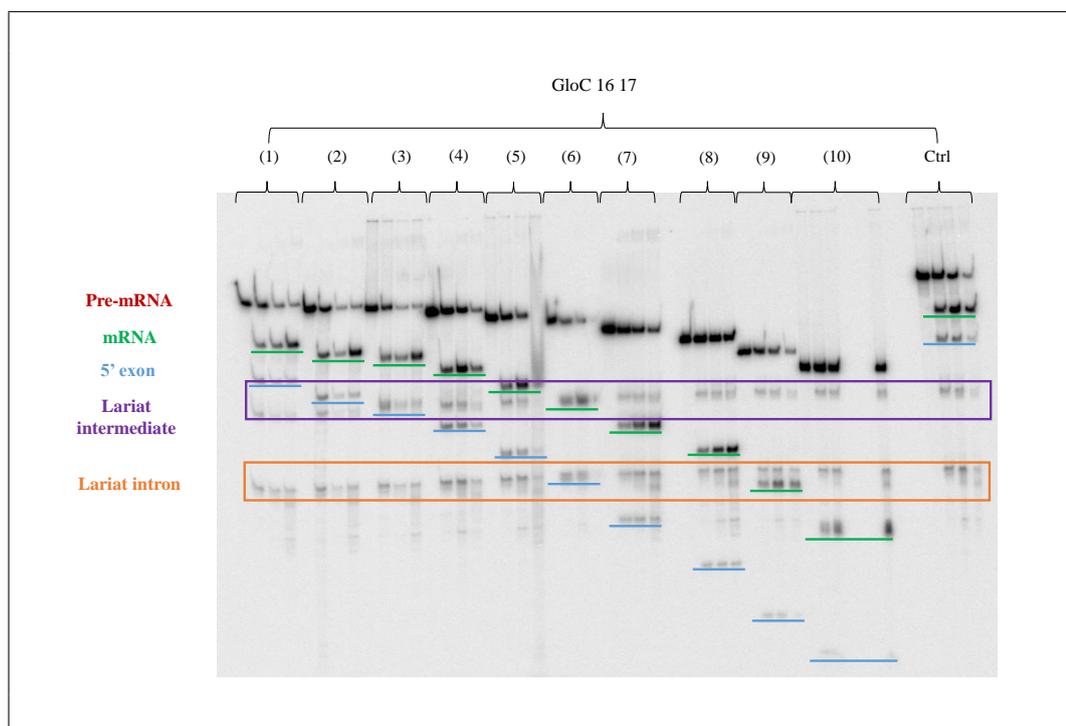


Figure 5.25: The *in vitro* splicing reaction for a series of GloC with different lengths. Ctrl represents the control experiments using a full length of GloC (labelled as GloC 16 17). Sample (1) indicates the next shorter length of GloC after the original one. Samples from (2) to (10) represent the shorter after sample number (1) and so on. Functional nuclear extract were use. 6% polyaccrelymide gel electrophoresis were implemented to analyse the splicing results. The gel was scanned via Typhoon. Number 17 represents the name of primer used in the PCR experiments.

## 5.5 SRm160 Vs. the 5' exon

---

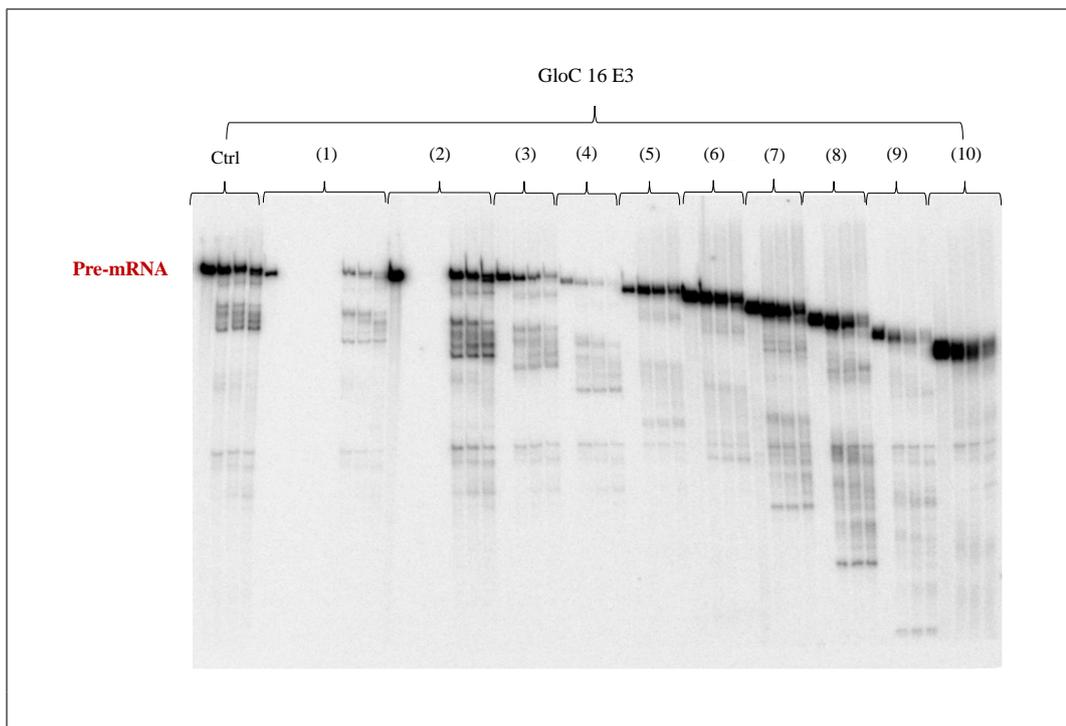


Figure 5.26: The *in vitro* splicing reaction for a series of GloC with different lengths. - Ctrl represents the control experiments using a full length of GloC but without the 3'-U1-binding site (labelled as GloC 16 E3). Sample (1) indicates the next shorter length of GloC after the original one. Samples from (2) to (10) represent the shorter after sample number (1) and so on. Functional nuclear extract were use. 6% polyaccrelymide gel electrophoresis were implemented to analyse the splicing results. The gel was scanned via Typhoon. E3 represents the name of primer used in the PCR experiments.

## 5.5 SRm160 Vs. the 5' exon

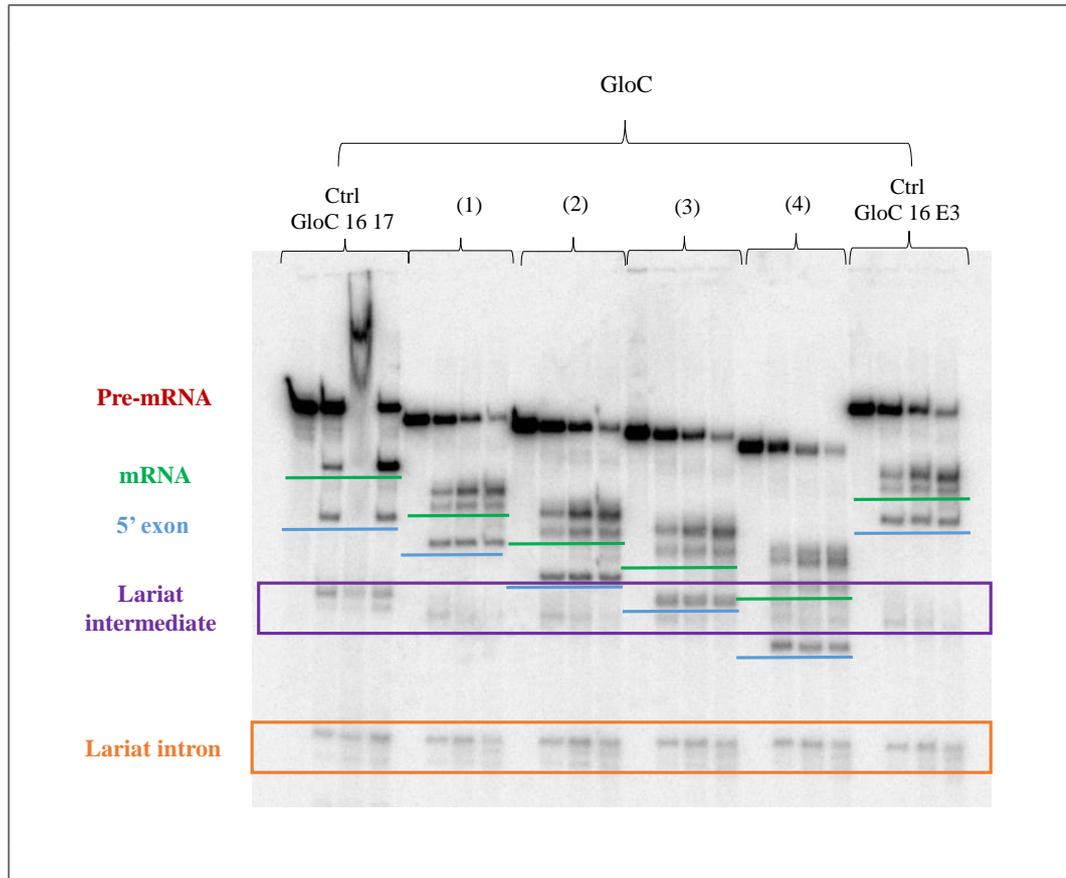


Figure 5.27: The *in vitro* splicing reaction for a series of GloC with different lengths. Two Ctrl were used; one represents the control experiments using a full length of GloC but with the 3'-U1-binding site (labelled as GloC 16 17) and the other represents the control experiments using a full length of GloC but without the 3'-U1-binding site (labelled as GloC 16 E3). Sample (1) indicates the next shorter length of GloC after the original one. Samples from (2) to (4) represent the shorter after sample number (1) and so on. Functional nuclear extract were use. 6% polyaccrelymide gel electrophoresis were implemented to analyse the splicing results. The gel was scanned via Typhoon. Numbers 16, 17 and E3 are the name of primers used in the PCR experiments.

## 5.5 SRm160 Vs. the 5' exon

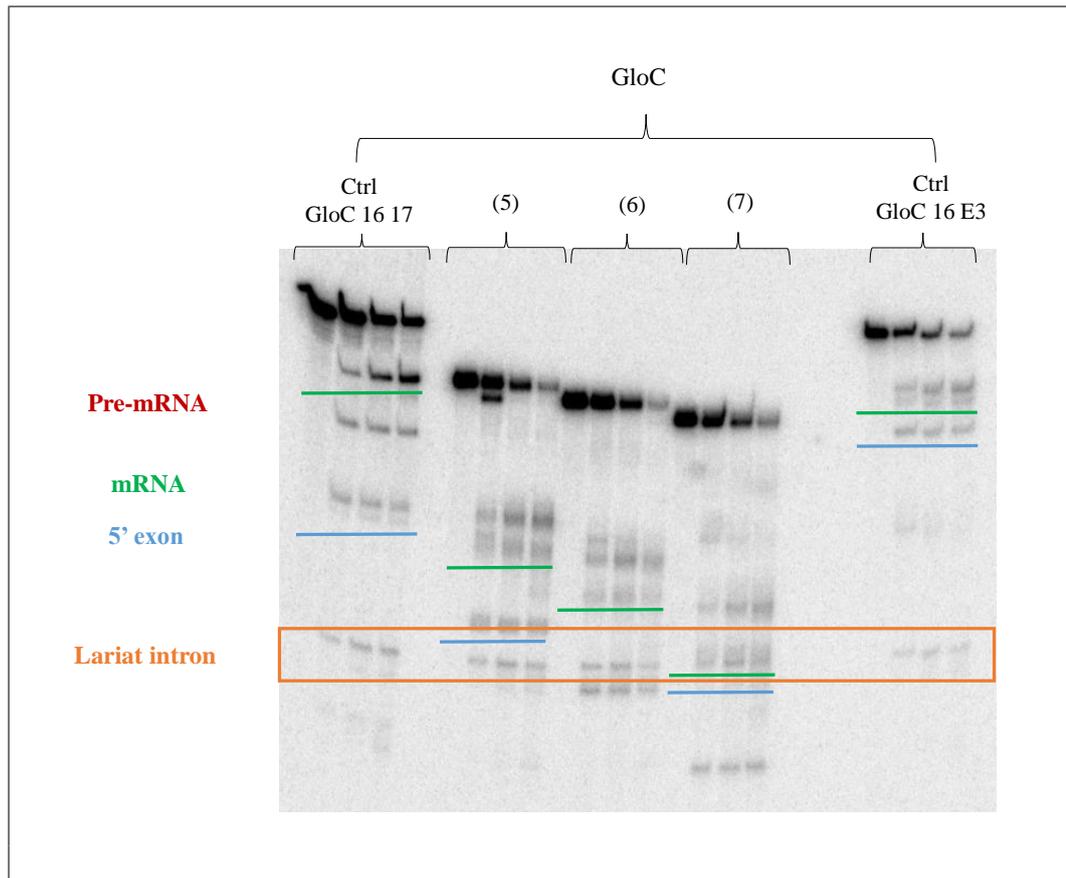


Figure 5.28: The *in vitro* splicing reaction for a series of GloC with different lengths. Two Ctrl were used; one represents the control experiments using a full length of GloC but with the 3'-U1-binding site (labelled as GloC 16 17) and the other represents the control experiments using a full length of GloC but without the 3'-U1-binding site (labelled as GloC 16 E3). Samples from (5) to (7) represent the shorter substrates after sample number (1) and so on. Functional nuclear extract were use. 6% polyaccrelymide gel electrophoresis were implemented to analyse the splicing results. The gel was scanned via Typhoon. Numbers 16, 17 and E3 are the name of primers used in the PCR experiments.



## 5.5 SRm160 Vs. the 5' exon

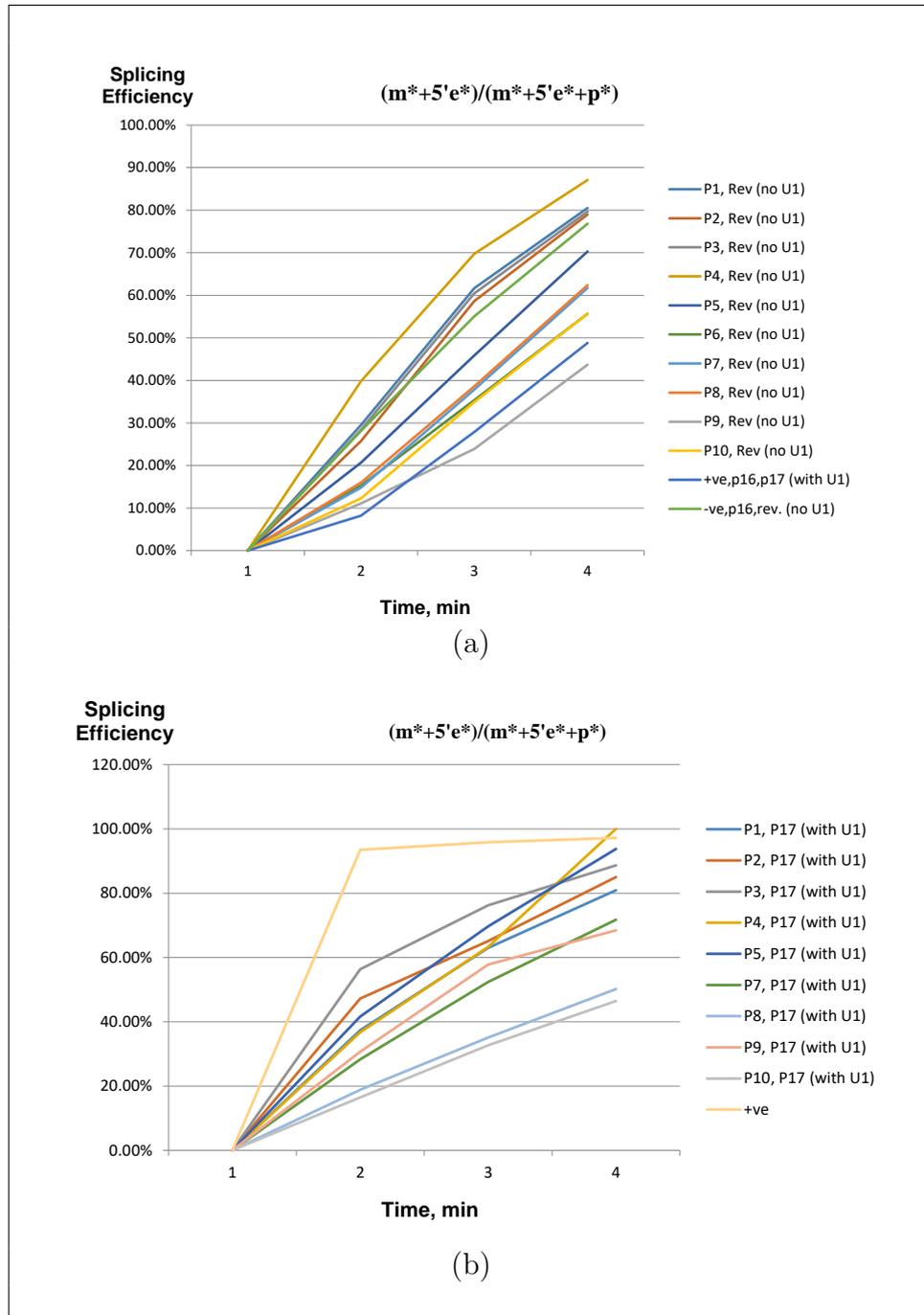


Figure 5.30: The *in vitro* splicing efficiency for a series of different length of GloC substrates. A comparison between substrates with and without the 3' U1 binding site. The two charts represent the splicing efficiency calculated depends on the mRNA product. The top one represents the results for all substrates without the 3' U1 binding site. The bottom one represents the results for all substrates with the 3' U1 binding site.

## 5.5 SRm160 Vs. the 5' exon

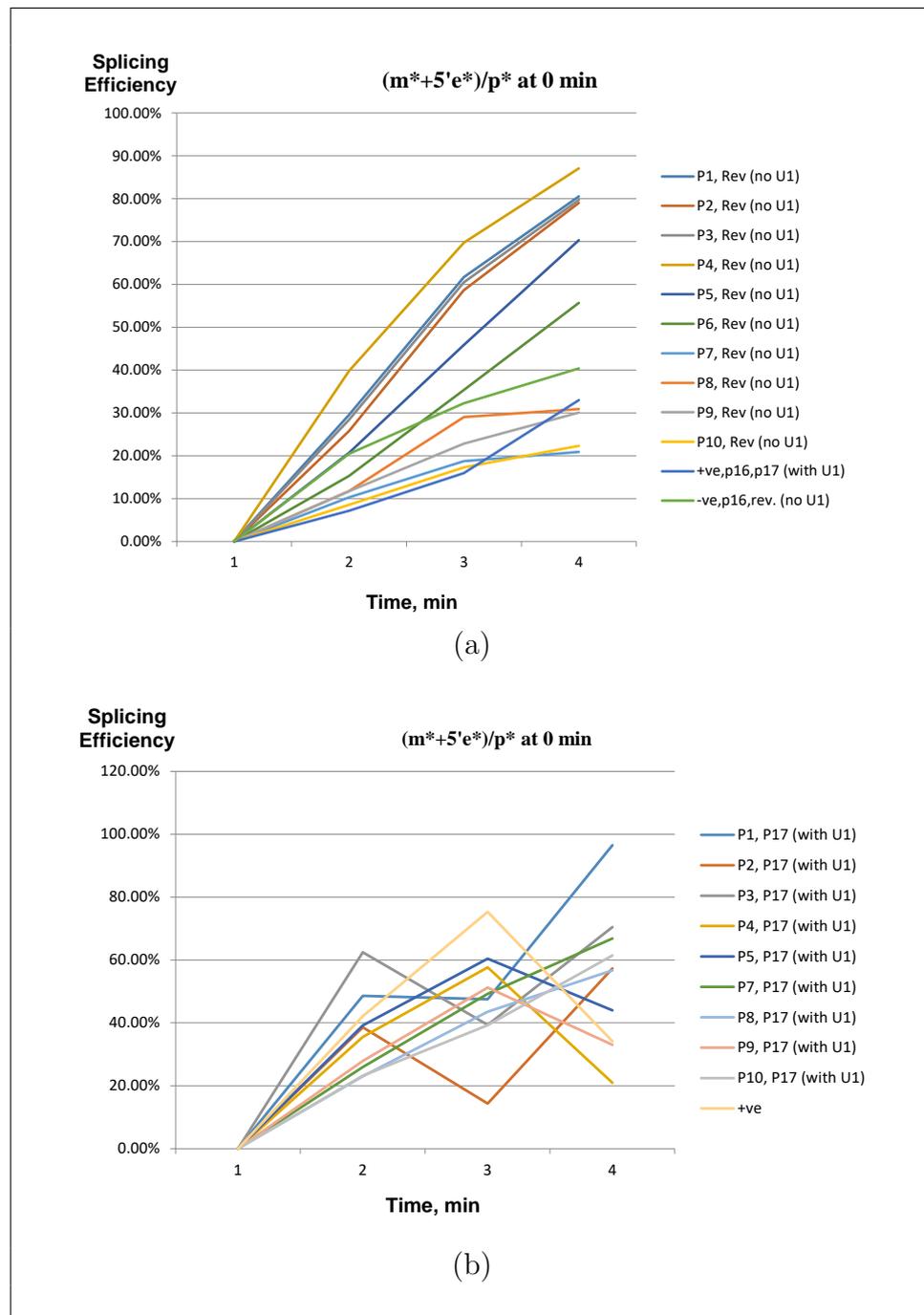


Figure 5.31: The *in vitro* splicing efficiency for a series of different length of GloC substrates. A comparison between substrates with and without the 3' U1 binding site. The two charts represent the splicing efficiency calculated depends on the pre-mRNA product at the zero time point. The top one represents the results for all substrates without the 3' U1 binding site. The bottom one represents the results for all substrates with the 3' U1 binding site.

## 5.5 SRm160 Vs. the 5' exon

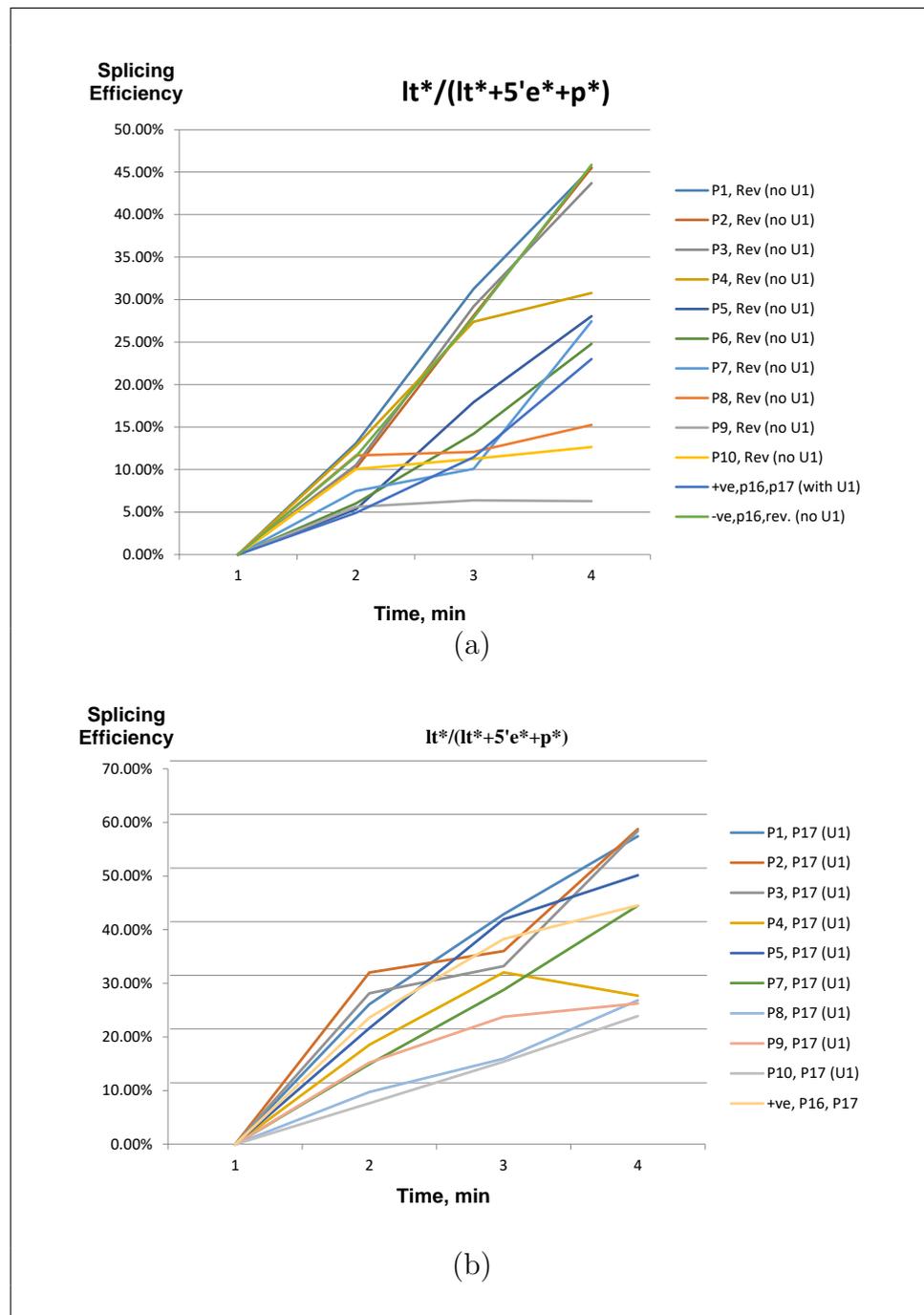


Figure 5.32: The *in vitro* splicing efficiency for a series of different length of GloC substrates. A comparison between substrates with and without the 3' U1 binding site. The two charts represent the splicing efficiency calculated depends on lariat product. The top one represents the results for all substrates without the 3' U1 binding site. The bottom one represents the results for all substrates with the 3' U1 binding site.

## 5.5 SRm160 Vs. the 5' exon

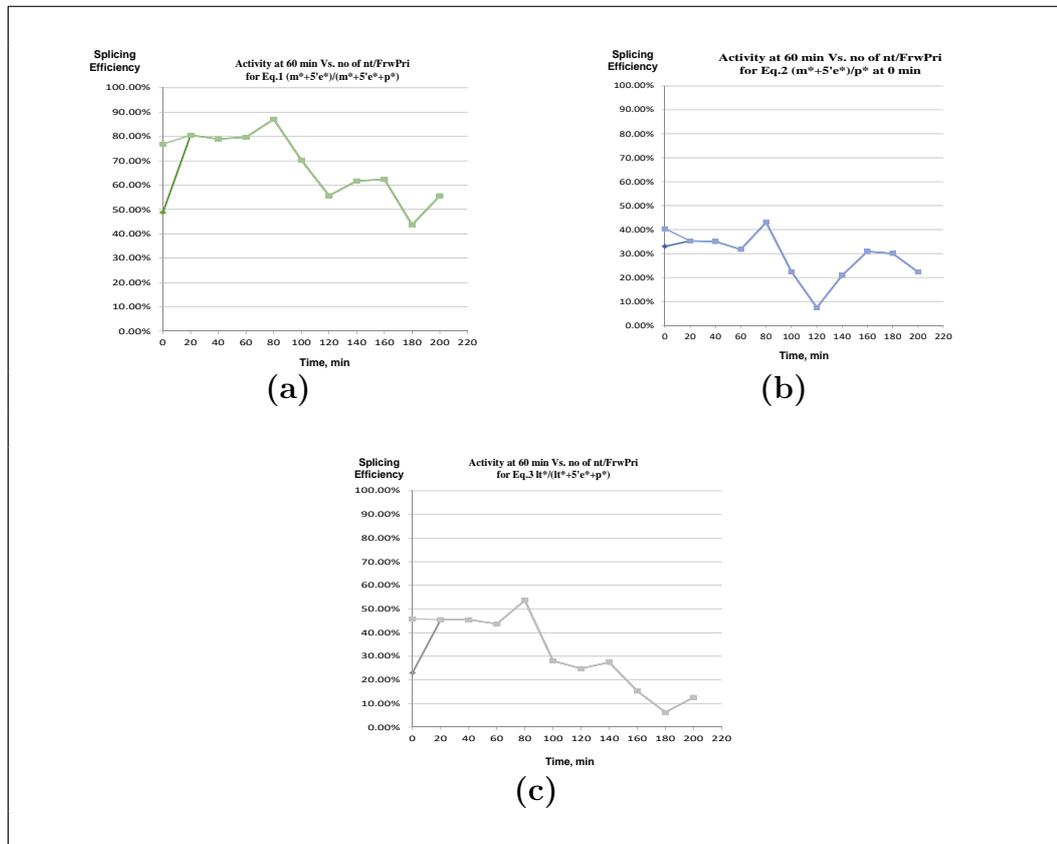


Figure 5.33: The *in vitro* splicing activity for a series of different length of GloC substrates. All substrates were without the 3' U1 binding site. **a** The activity was calculated for all substrates depending on the mRNA results. **b** The activity was calculated for all substrates depending on the pre-mRNA results at the zero time point. **c** The activity was calculated for all substrates depending on the lariat intron results.

## 5.5 SRm160 Vs. the 5' exon

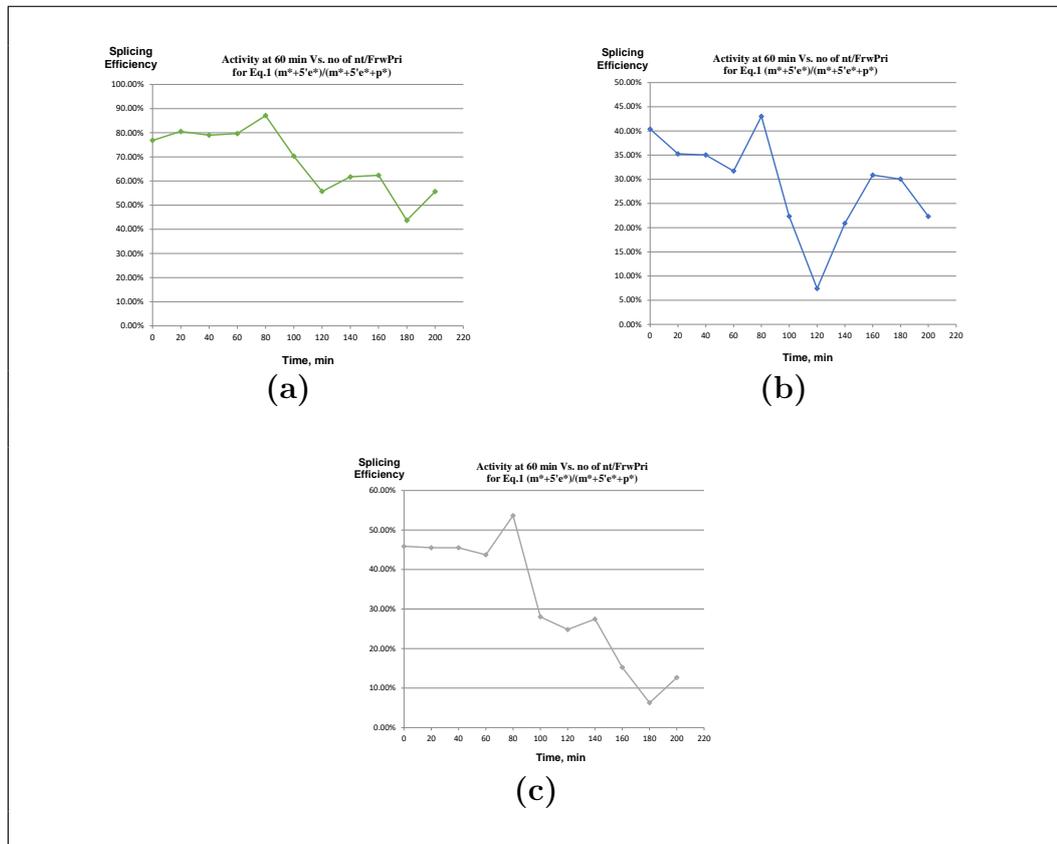


Figure 5.34: The *in vitro* splicing activity for a series of different length of GloC substrates. All substrates were with the 3' U1 binding site. **a** The activity was calculated for all substrates depending on the mRNA results. **b** The activity was calculated for all substrates depending on the pre-mRNA results at the zero time point. **c** The activity was calculated for all substrates depending on the lariat intron results.

## 5.5 SRm160 Vs. the 5' exon

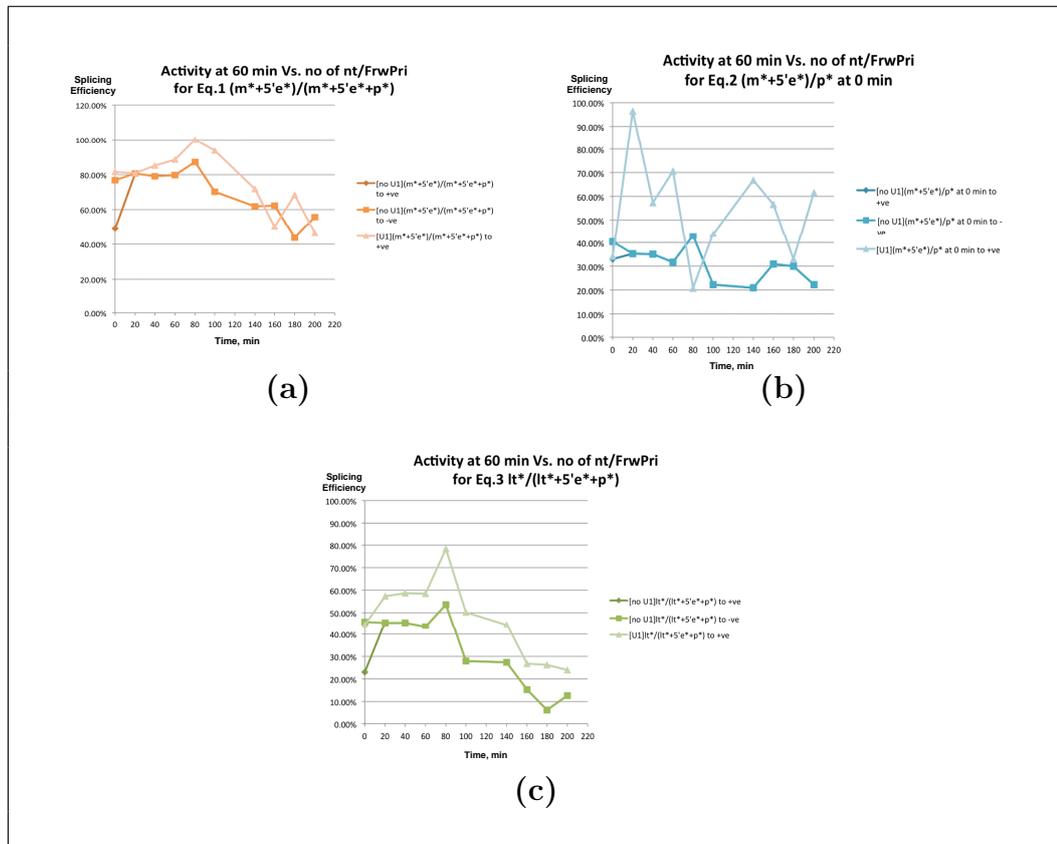


Figure 5.35: The *in vitro* splicing activity for a series of different length of GloC substrates. A comparison between substrates with and without the 3' U1 binding site. **a** The activity was calculated for all substrates depending on the mRNA results. **b** The activity was calculated for all substrates depending on the pre-mRNA results at the zero time point. **c** The activity was calculated for all substrates depending on the lariat intron results.

## Chapter 6

# Analysis of Complexes Flexibility Using FRET

This chapter begins with a description of the donor and acceptor dyes used to demonstrate the FRET interaction and why these dyes were chosen. The modification of the sample chamber is also discussed. Finally, the general experimental procedures and characterisation techniques are presented.

# Contents

6.1	Fluorescence Dyes . . . . .	175
6.2	Fluorescence Resonance Energy Transfer; FRET . . . . .	175
6.2.1	Branched Substrate . . . . .	176
6.2.2	Double-labeled RNA . . . . .	187

### 6.1 Fluorescence Dyes

In the first part of the work presented here, FRET was achieved using a donor dye called ATTO. Its fluorescence emission band is at 610 nm. The acceptor is Cy5 that fluorescence at 650 nm. The donor ATTO overlaps sufficiently with the absorption band of the Cy5 acceptor dye. The ATTO dye has many benefits including its strong absorption, high quantum efficiency, and large Stokes shift. Cy5 is also one of the most widely used fluorescent dyes for labelling biological molecules. The absorption and the emission wavelengths are in the visible red spectral region, which is sufficiently removed from the intrinsic fluorescence of most biological tissues, to suppress the background noise. As a FRET pair, those two dyes have a clear Stokes shift, which reduces cross talk and improves the accuracy of the results.

One of the drawbacks to these dyes is that they are photosensitive and can photo-bleach easily when irradiated continuously, making them harder to handle. Photobleaching results in the loss of the dyes ability to absorb light of a particular wavelength, due to photo degradation of the dye molecules, thus reducing their emission. This can be minimized by keeping the dye samples in the dark and limiting their exposure to ambient light. Also, by using reagents to prevent bleaching such as scavengers.

### 6.2 Fluorescence Resonance Energy Transfer; FRET

Alongside single molecule colocalisation measurements, single molecule FRET microscopy was initiated in the start of a project to look at FRET signals under conditions supporting the formation of the early complex E, complex

## 6.2 Fluorescence Resonance Energy Transfer; FRET

---

A and the presence of anti-U1 oligo. Additionally, using our home-made extract (SRm160) might open a gate to more investigation of its state as a rigid molecule (160 KDa).

As previously explained in (Chapter 1), this technique was used to probe inter- or intra-molecular energy transfer between two fluorophores (donor and acceptor) on the range of 1 - 10 nm length scale making it ideal to investigate the structure, dynamics, and orientations of the molecule of interest.

Here in the following sections, few FRET results were achieved using our home-made single molecule microscope under different conditions.

### 6.2.1 Branched Substrate

Branched DNA structures (Figure 6.1) play critical roles in DNA replication, repair and recombination in addition to being key building blocks in DNA nanotechnology. This can be investigated in solution by detection of the internal distances between two fluorescent dyes, donor (D) and acceptor (A) via FRET. This structure is used here in this work as a simple model in order to optimize the FRET system.

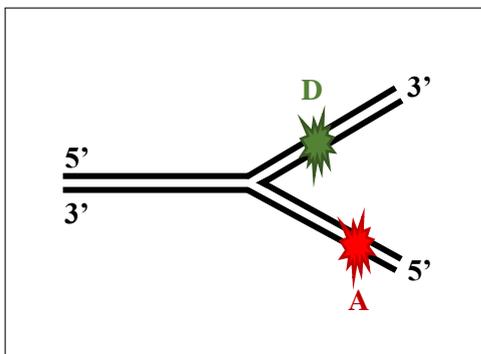


Figure 6.1: Schematic diagram of the DNA branched structure (forked DNA). - The solid black line represents two complementary DNA. - Star in green represents the donor fluorophore (D; Cy3). - Star in red represents the acceptor (A; Cy5).

## 6.2 Fluorescence Resonance Energy Transfer; FRET

---

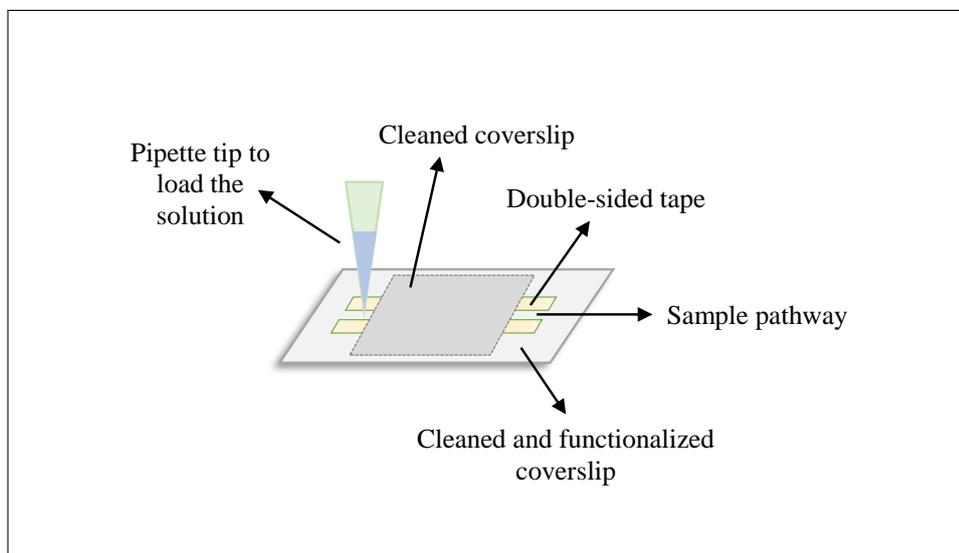


Figure 6.2: An exemplary of sample chamber. The chamber is prepared by sandwiching two cleansed coverslips by double-sided tape. The double-sided tape is used to create a channel in abut 2 mm. The sample is introduced into the chamber via pipette tip that contains the solution (sample of interest).

To perform FRET experiments, a slide chamber is designed by sandwiching double-sided tape between a pre-cleaned slide and a cover slip (Figure 6.2). For stable and specific immobilization of the sample to the slide, a biotin streptavidin linkage is used. Biotinylated bovine serum albumin (BSA) that adsorbs to slide surface is used. BSA binds the biotinylated molecules through streptavidin (Figure 6.3). The assembled chamber is checked for purities before and after application of BSA/Biotin or streptavidin by imaging the surface in the absence of any solution. If there are no fluorescent spots appeared, the cleansed slide is deemed acceptable. A biotinylated biomolecule is added in concentrations (nanometers) that allow us to achieve immobilization at the desired single-molecule density. Additionally, slides were tested after the addition of streptavidin and BSA in order to ensure there is no contamination can contribute with the final results (Figure 6.4)

## 6.2 Fluorescence Resonance Energy Transfer; FRET

---

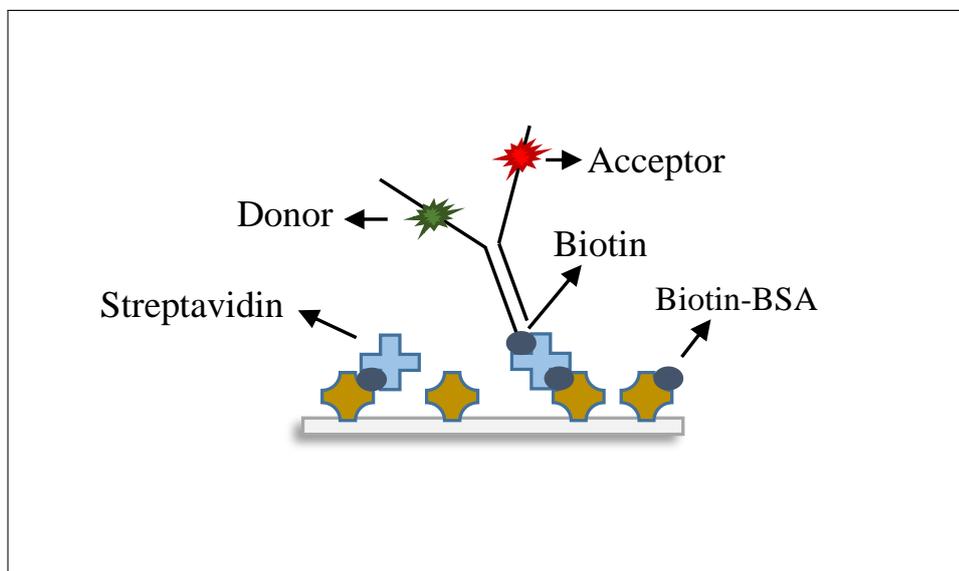


Figure 6.3: Sample immobilization. The cleaned slide is first treated via Biotin-BSA and then via streptavidin in order to minimize the non-specific binding to the slide surface. The sample of interest that has biotin at one end then introduced which can immobilize to the surface via streptavidin. The purity of each step was tested (see Figure 6.4).

In addition, cross-talk between the detected channels needs to be corrected before calculating the true FRET efficiency. To do this, sample of donor only and acceptor only are used. With sample contains donor-only, molecules were excited at a donor excitation wave lengths, the leakage of donor emission in to the acceptor channel is determined. Similar to donor-only, the leakage of acceptor emission via its excitation wave length into donor channel is also identified. The obtained results indicate no cross-talk is measured. Additionally, the cross-talk was optimized using specific correction filters suitable for both donor (ATTO647N) and Acceptor (Cy5), see (Figure 6.5). Those figures were a result of experiments done in the absence of any correction filters. The obtained results shows appearance of a signal on the acceptor detection panel which refer to cross-talk. The cross-signal could contribute with the real

## 6.2 Fluorescence Resonance Energy Transfer; FRET

---

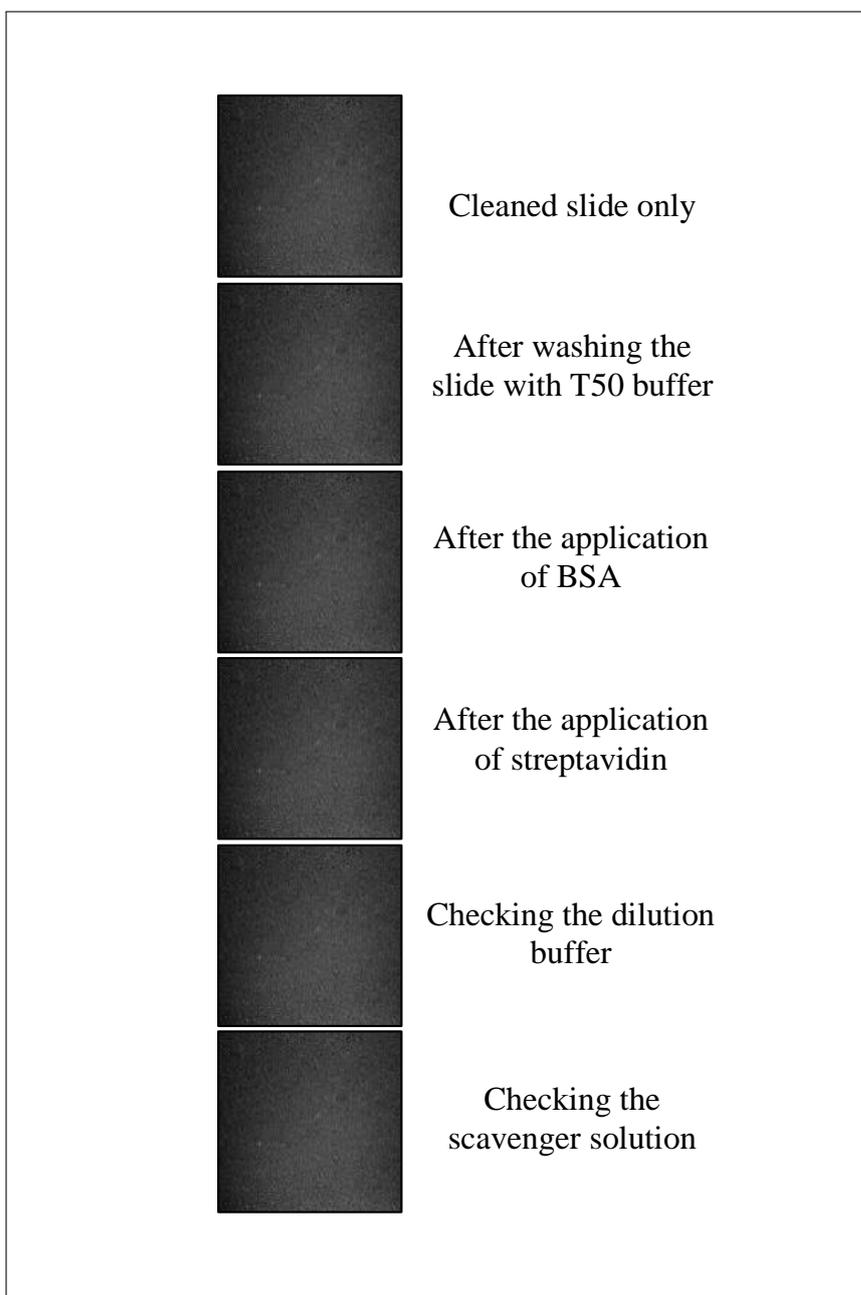


Figure 6.4: Testing cover slips (slide champer) before and after the addition of any solutions (streptavidin and BSA). The obtained results identified that all reagents (streptavidin, BSA and imaging buffers) are acceptable for single molecule FRET experiments. Each image is labelled via its condition on the right side. All the obtained results were measured using single molecule FRET microscope.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

data. To solve this, correction filters specific for Cy3 and (ATTO647N) were used. The resulted signals shows no cross-talk is exist, see (Figure 6.6). Moreover, a fluorescent microspheres sample is used in order to ensure an accurate correspondence between the donor and acceptor images.

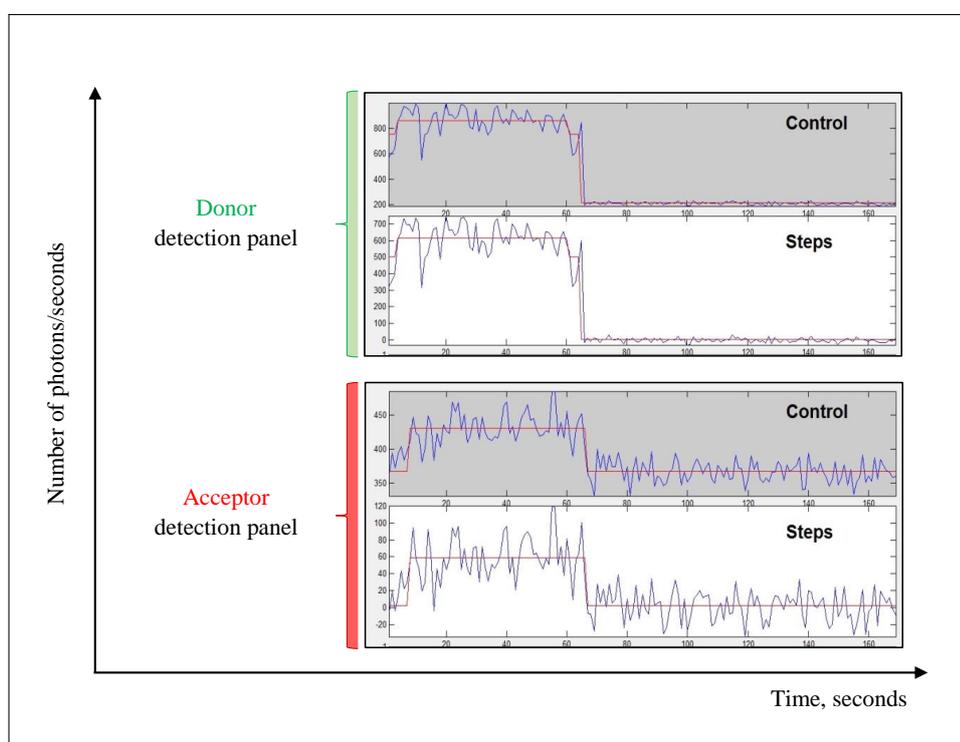


Figure 6.5: Schematic diagram of Cy3-Sample only. Single molecule experiments were implemented at 532 nm. No correction filters were used. Slides were treated via BSA and Streptavidin. The X axis represents the intensity. The Y axis represents the time in seconds. - The top panel represents the donor detection panel. The bottom panel represents the donor detection panel.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

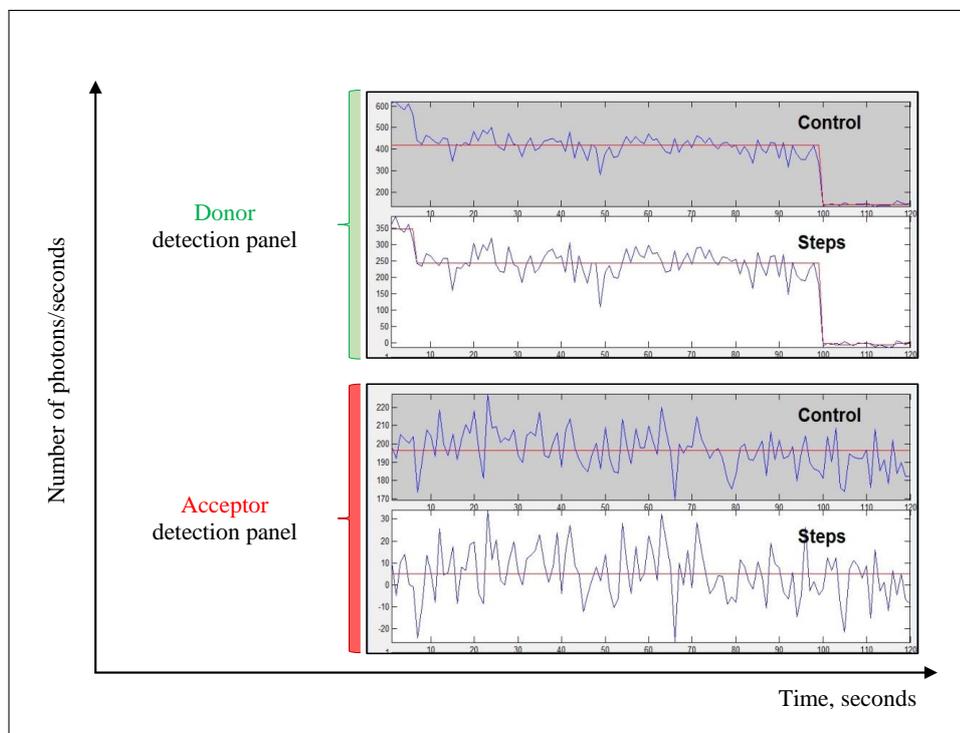


Figure 6.6: Schematic diagram of Cy3-Sample only. Single molecule experiments were implemented at 532 nm. Correction filters were used - Slides were treated via BSA and Streptavidin. The X axis represents the intensity. The Y axis represents the time in seconds. - The top panel represents the donor detection panel. The bottom panel represents the donor detection panel.

For FRET measurements, a biotinylated branched molecule (DNA) is used which has a single donor dye (Cy3) and an acceptor dye (Cy5), see figure 6.1. Immobilization was done via BSA-streptavidin linkage as described previously.

FRET signals were acquired by imaging surface immobilized molecules with the objective-TIRF microscope. An oxygen scavenger, such as PCA (Protocatechuic Acid) and PCD (Protocatechuate-3,4-dioxygenase) were used in combination with Ascorbic acid and Parquet to suppress the photobleach and the fluorescence instability, respectively.

At the beginning of FRET project, a prism-based TIRF microscope was

## 6.2 Fluorescence Resonance Energy Transfer; FRET

---

used. The microscope was optimised via Dr. Dmitry Cherny. Experiments were modified and compared in order to optimise conditions. The obtained results show a FRET signal (Figure 6.7 and 6.8). The single molecule FRET time trajectories from the branched DNA (forked DNA) indicates two bleaching steps. This might refer to two conformational states that exist.

Moreover, similar experiments were performed via the objective-based TIRF microscope. The one used with single molecule colocalisation experiments. All conditions were repeated and modified to be suitable under the objective-based microscope. The obtained results were similar. FRET signals were recorded. The analysed bleaching steps indicate two bleaching steps. This results confirm the findings via prism-based TIRF microscope (Figure 6.9 and 6.10)

## 6.2 Fluorescence Resonance Energy Transfer; FRET

---

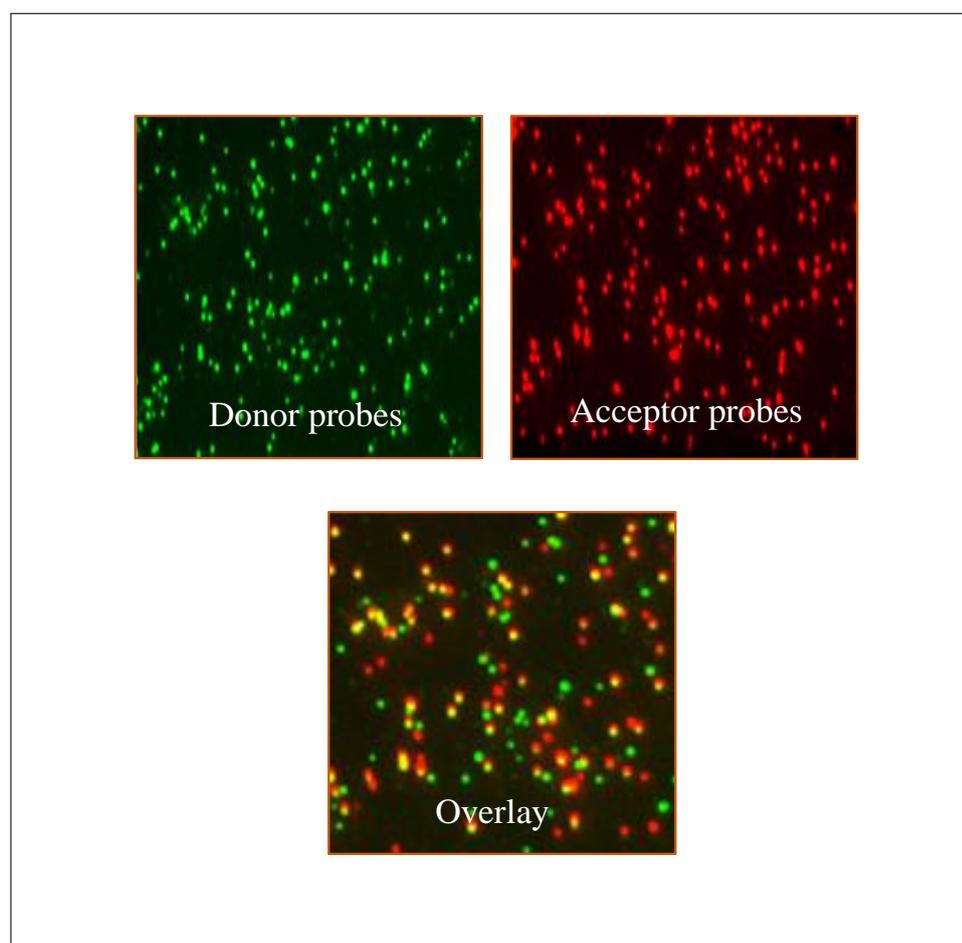


Figure 6.7: Total internal reflection fluorescence microscopy imaging. View from fluorescence microscopy, Prism-based TIRF is used to look at the FRET of the forked DNA. Top row; donor probes which is represented as green spots and acceptor probes which appears as red spots. Bottom row; the overlay image showing colocalized spots which are represented as yellow spots. Green spot refers to Cy3 fluorophores. Red spots refer to Cy5 fluorophores.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

---

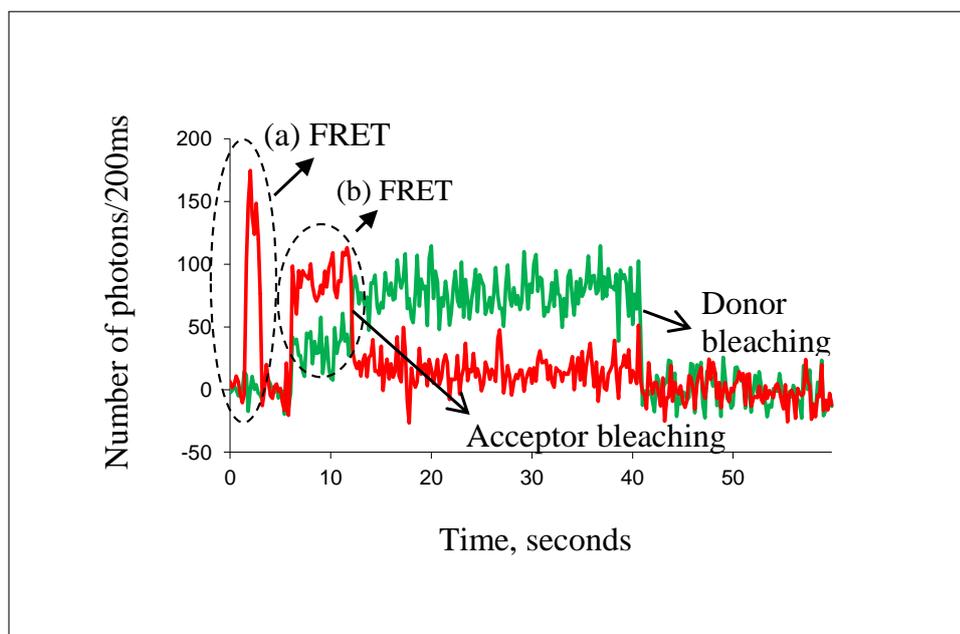


Figure 6.8: Total internal reflection fluorescence microscopy imaging. Representative single molecule FRET time trajectories from the branched DNA. Green and red traces represent the fluorescence intensity of the donor dye (CY3) and the acceptor dye (CY5), respectively. **(a)** High FRET signal. **(B)** Second state showing FRET.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

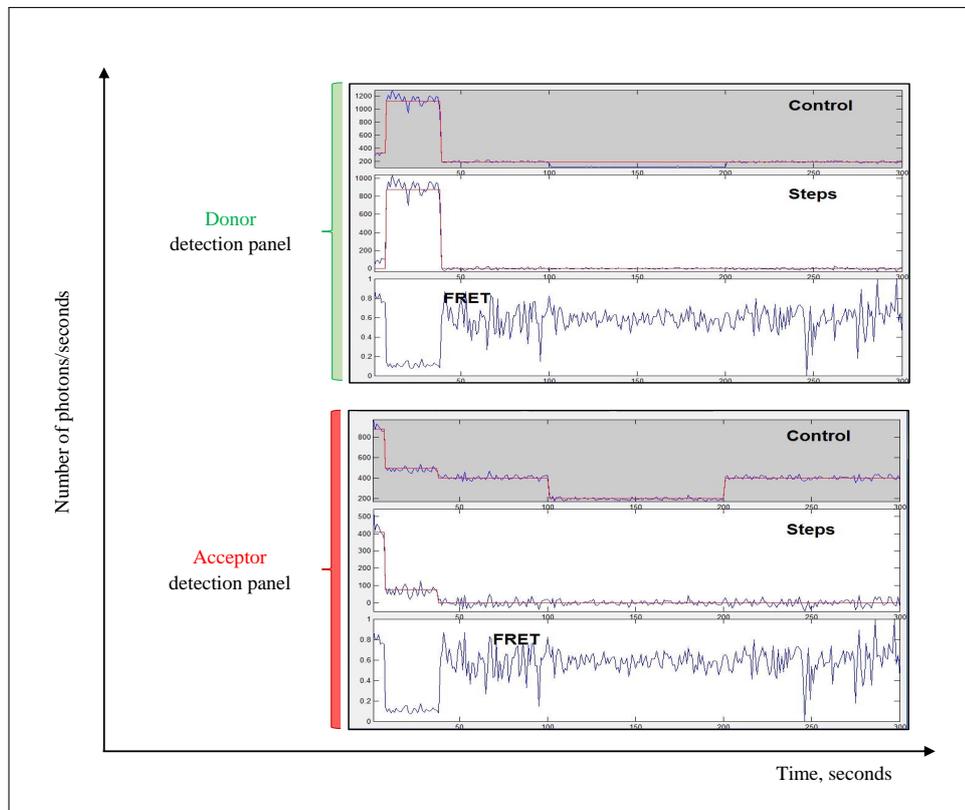


Figure 6.9: Schematic diagram of FRET signal of Forked DNA sample. Single molecule FRET experiments were implemented under the condition supported the optimization of FRET). Treated Slide chamber were used.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

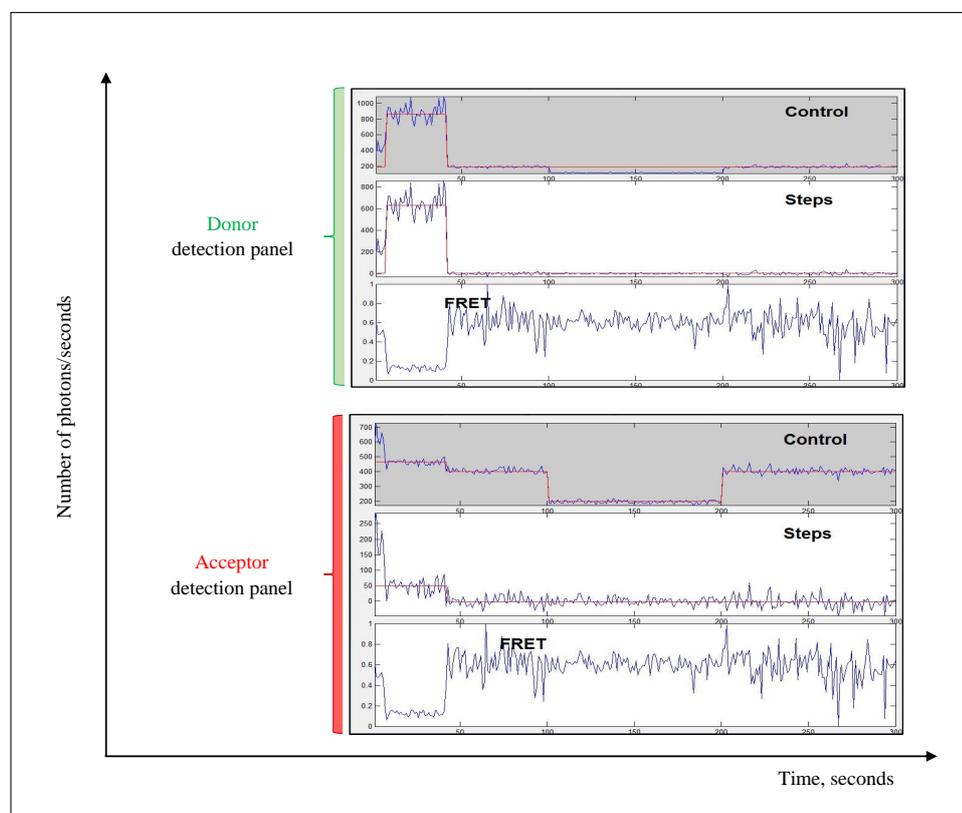


Figure 6.10: Schematic diagram of FRET signal of Forked DNA sample. Single molecule FRET experiments were implemented under the condition supported the optimization of FRET). Treated Slide chamber were used.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

---

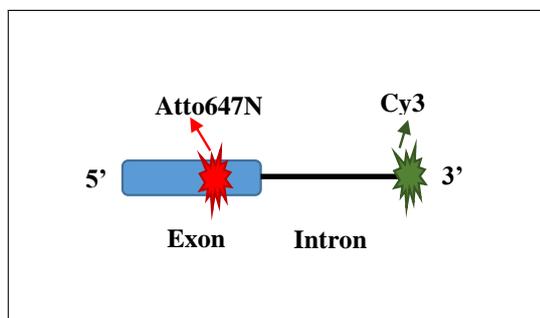


Figure 6.11: The structure of the double-labelled RNA construct. From left; the blue box represents the 5' exon, star in red represents the acceptor dye (ATTO647N), solid line represents the intron and the star in green represents the donor fluorophore (Cy3) which incorporate to the RNA at its 3' end.

### 6.2.2 Double-labeled RNA

Single molecule FRET was used to study the effect of nuclear extract media in the absence and presence of ATP and snRNPs such as U1 snRNP on the conformational changes and dynamics of RNA of interest.

To do this experiment, FRET system was optimized as previously explained. Pre-cleaned slides were modified and treated via BSA/Biotin and streptavidin. Simple biotinylated double-labeled RNA substrate was used (a kind gift from Prof.Eperon). This RNA consists of a 5' exon containing ESE sequence, a bit of intron (about 13 nucleotides), a single donor dye (Cy3) at the 3' end and an acceptor dye (ATTO647N) at 5' exon (Figure 6.11). Functional commercial nuclear extract was available. Home made nuclear extract (SRm160) was also introduced. Samples were prepared in the absence and presence of ATP, diluted and injected to the modified slides. An oxygen scavenging system was used too.

With the double-labeled RNA, an appropriate wavelength (532nm) was used in order to excite the donor molecule (ATTO647N) and to asses FRET experiments. In this experiment, FRET is expected to identify when the

## 6.2 Fluorescence Resonance Energy Transfer; FRET

---

ATTO647 and Cy3 are brought close together. In the presence of ATP, FRET signals between ATTO647N and Cy3 were monitored over time (Figure 6.14 and Figure 6.15). The FRET time traces results of individual molecules over time do not exhibit significant fluctuations, indicating that the observed FRET arises from relatively static heterogeneity among RNA rather than conformational fluctuations. This static FRET is probably because of the presence of rigid molecule (SRm160; 160kDa).

The absence of ATP revealed no FRET to be made between ATTO647N and Cy3 dyes (Figure 6.12 and Figure 6.13).

In contrast, signals were observed in the ATTO647N detected channel after 50 seconds, presumably because molecules are getting out of focus during the experiment. Interestingly, addition of an oligonucleotide complementary to U1 snRNP in the presence of ATP showed quite similar (Figure 6.16). This result suggest two possibilities; very low FRET identified between the two dyes, or unknown transition is occurred.

All previous FRET data were analysed via Dr. Robert Weinmeister's program which originally designed for colocalisation analysis [Weinmeister, 2014]. In order to detects spots and hence calculate the FRET, multiple steps were suggested by Dr. Robert Weinmeister as following

- a. Manually click on a spot from a donor channel, search and click on the corresponding spot on the acceptor channel.
- b. Repeat the previous step for each spot.
- c. Calculate the traces.
- d. Export data the Excel file.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

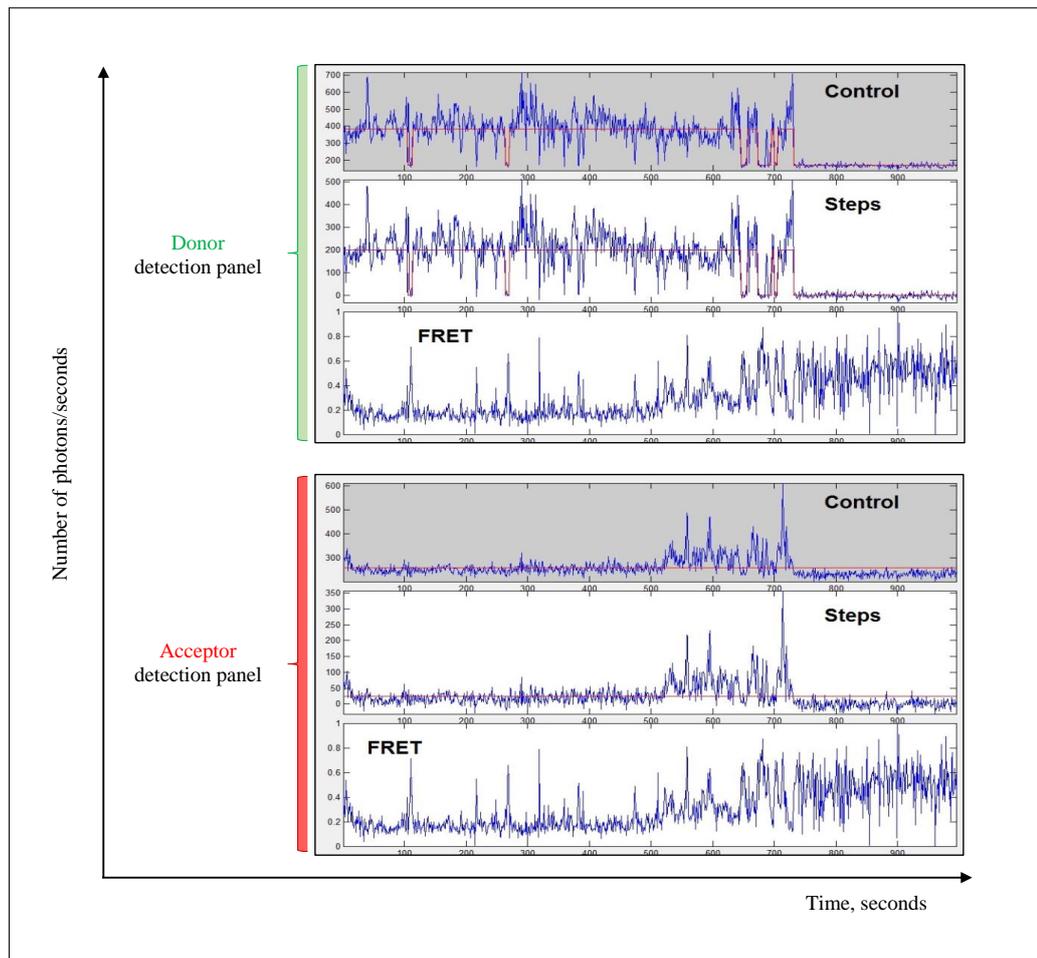


Figure 6.12: Schematic diagram of FRET signal of the double-labelled RNA. Single molecule FRET experiments were implemented under the condition supported the formation of the early complex E (-ATP). Functional home made nuclear extract (SRm160) was used.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

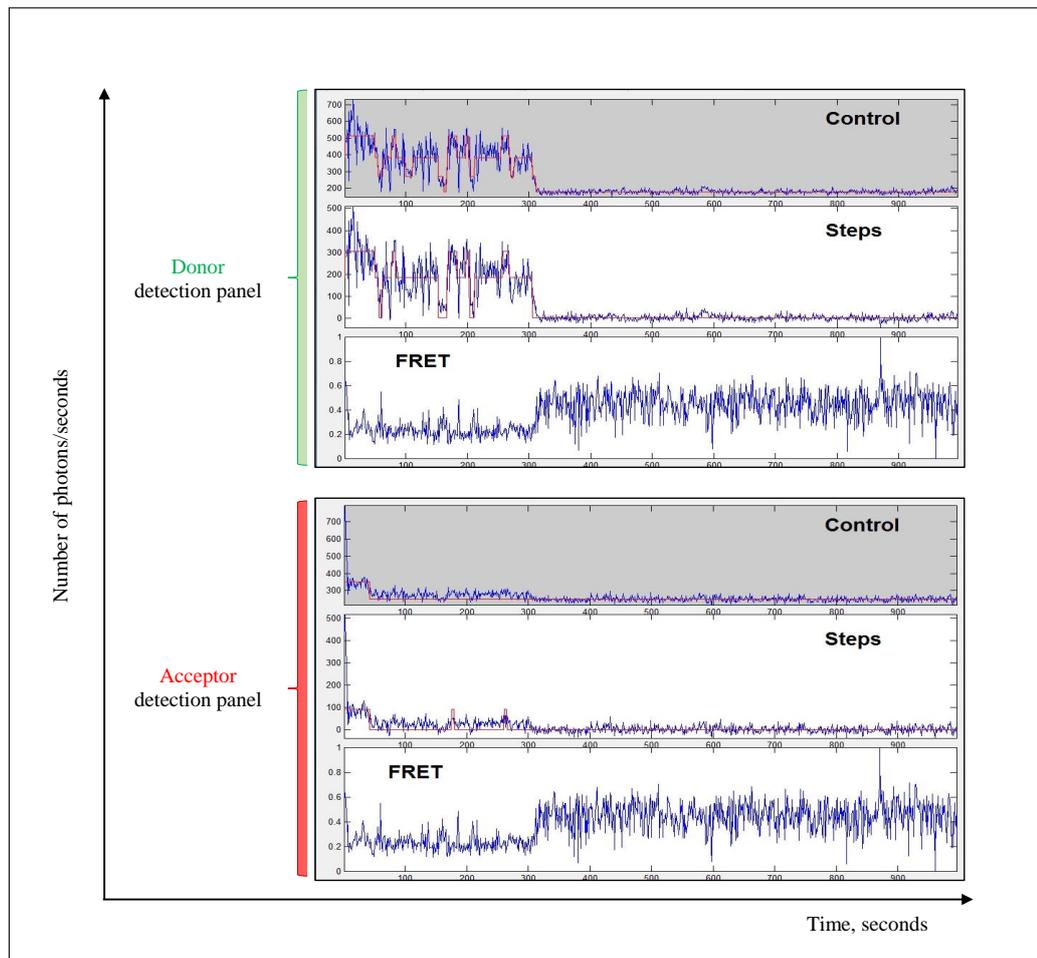


Figure 6.13: Schematic diagram of FRET signal of the double-labelled RNA. Single molecule FRET experiments were implemented under the condition supported the formation of the early complex E (-ATP). Functional home made nuclear extract (SRm160) was used.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

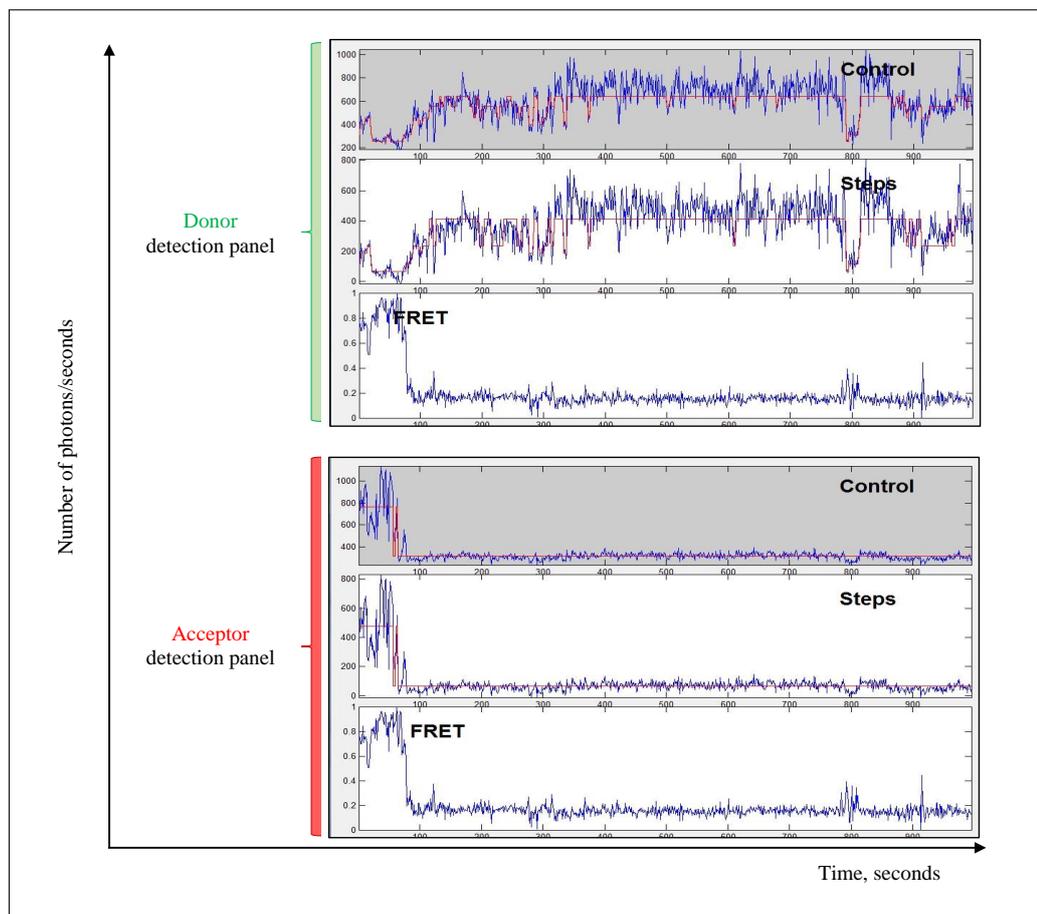


Figure 6.14: Schematic diagram of FRET signal of the double-labelled RNA. Single molecule FRET experiments were implemented under the condition supported the formation of complex A (+ATP) in the presence of anti-U6 oligo). Functional home made nuclear extract (SRm160) was used.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

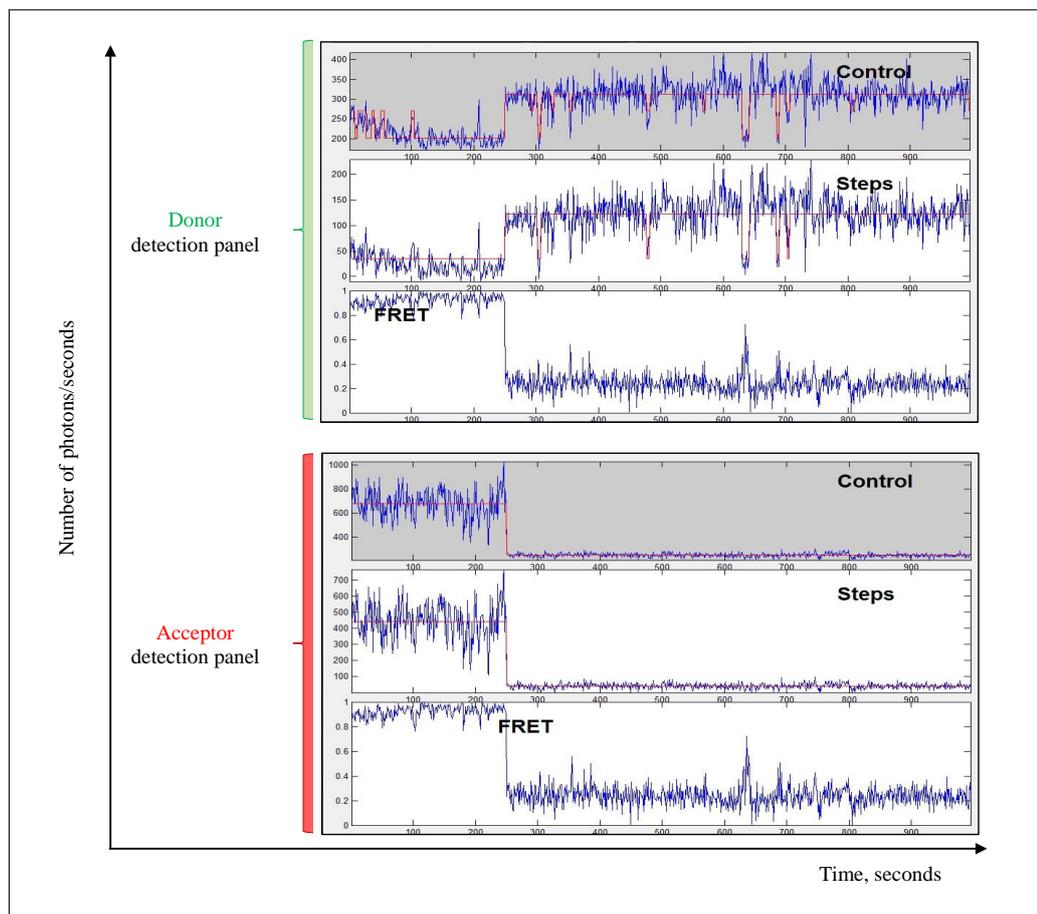


Figure 6.15: Schematic diagram of FRET signal of the double-labelled RNA. Single molecule FRET experiments were implemented under the condition supported the formation of complex A (+ATP) in the presence of anti-U6 oligo). Functional home made nuclear extract (SRm160) was used.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

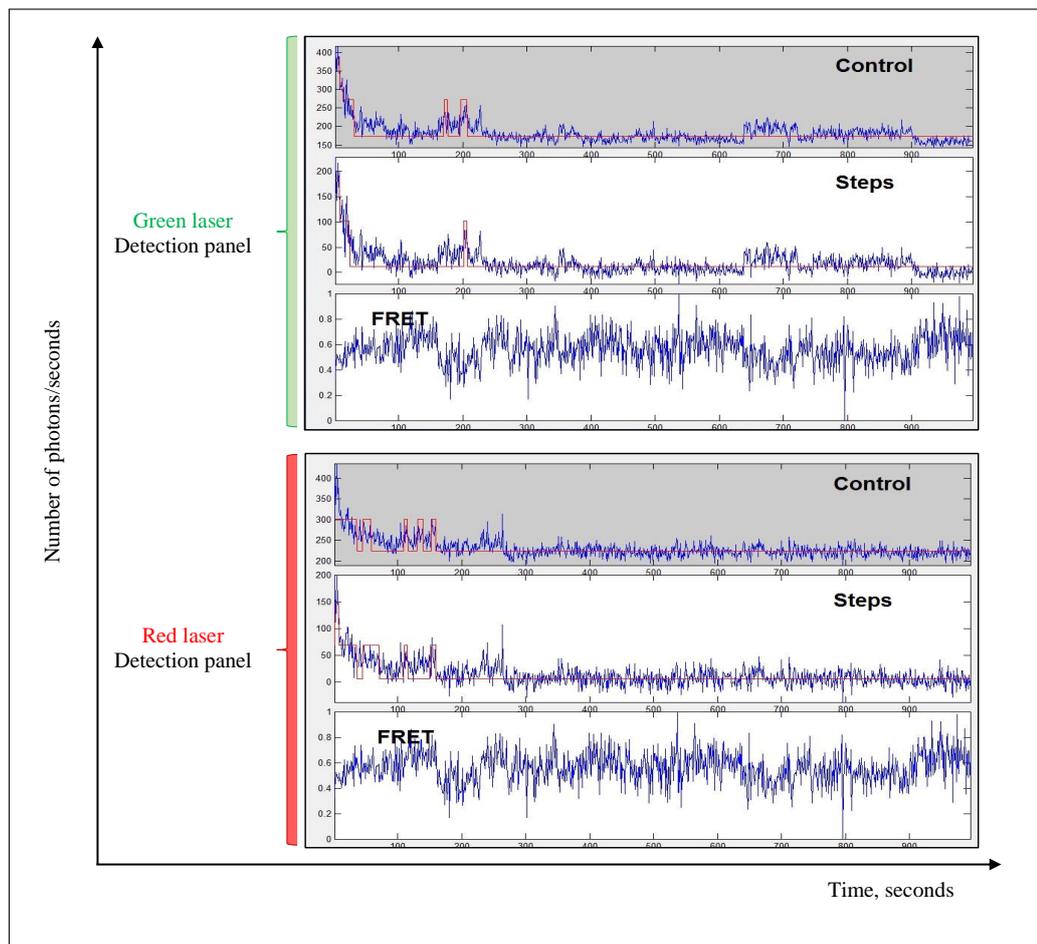


Figure 6.16: Schematic diagram of FRET signal of the double-labelled RNA. Single molecule FRET experiments were implemented under the condition supported the formation of complex A (+ATP) in the presence of anti-U1 oligo. Functional home made nuclear extract (SRm160) was used.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

---

However, those steps have to be done manually. As each spot has two traces, one corresponds to donor channel and the other corresponds to acceptor channel, a screen shot has to be taken for each trace panel in order to be represented in a document. All traces were plotted in one color and the overlap signal does not exist which makes the comparison difficult between the donor and the acceptor. Also, if the number of detected spots on one channel are not equal to those on the corresponding channel, interfering occurs in the total number of traces. These steps are subject to multiple errors and time consuming.

To overcome these limitations I developed a tool by Matlab with help from Dr. Alaa Khadidos. The tool reads the corresponding cells from each Excel file automatically to calculate the signal of the donor spot (in green), the acceptor spot (in red), the overlap signal between them and the FRET signal through different number of frames. Automatically, all resulted data are stored in different formats such as .EPS and .PDF to easily be represented in a document.

Moreover, we developed a program that allows us to automatically calculate the FRET signals by clicking spots on donor channel. The corresponding spots on the acceptor channel are automatically detected, analysed and stored. The Figure 6.17 shows an example of a spot calculated using ImageJ, Robert Weinmeister's and our program. Note that our program is still under development.

## 6.2 Fluorescence Resonance Energy Transfer; FRET

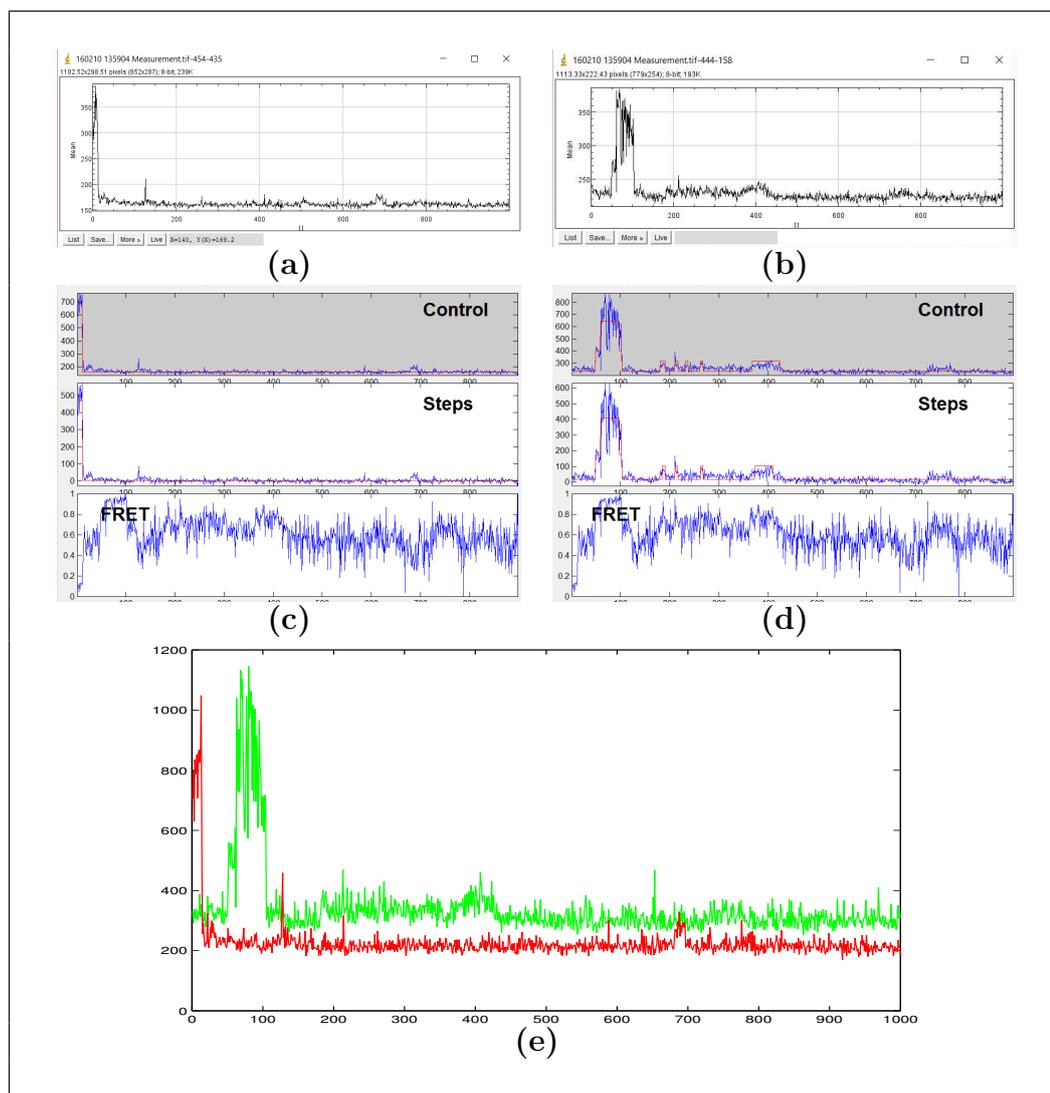


Figure 6.17: A comparison diagrams for single molecule signals through different programs. (a) A time trace for the acceptor fluorophore was analysed via ImajJ. (b) A time trace for the donor fluorophore was analysed via ImajJ. (c) A time trace for the acceptor fluorophore was analysed via Dr. Robert program. (d) A time trace for the acceptor fluorophore was analysed via Dr. Robert program. (e) A time trace shows the overlap between the acceptor fluorophore and the donor fluorophores analysed via our additional program. Same spots were detected in order to compare.

# Chapter 7

## Discussion

The main purpose of this research was to investigate whether the co-activator splicing factor (SRm160) is involved in: the early complex E or complex A, influenced by a specific RNA construct, depends on the presence of ESE sequences, recognizing candidate sites or whether it binds to the selected sites in response to other factors. In addition, it is important to investigate whether SRm160 is acting as a large single molecule (as a bridge) binding to 5' splice site through the U1 snRNP, to the branch site through the U2 snRNP, and to the 3' exon via other SR proteins at the same time as it was believed in all previous researches.

# Contents

7.1	The Microscope . . . . .	202
7.2	The Analysis . . . . .	206
7.3	The Preparation . . . . .	208
7.4	The Protein; SRm160 . . . . .	209
	7.4.1 The SM-Colocalisation Vs. SRm160 . . . . .	209
	7.4.2 The FRET signal Vs. SRm160 . . . . .	216
7.5	Summary . . . . .	217
7.6	Future Work . . . . .	221

## CONTENTS

---

*Reasoning draws a conclusion,  
but does not make the conclusion  
certain, unless the mind discusses  
it by the path of experience ....*

*Roger Bacon ..*

## CONTENTS

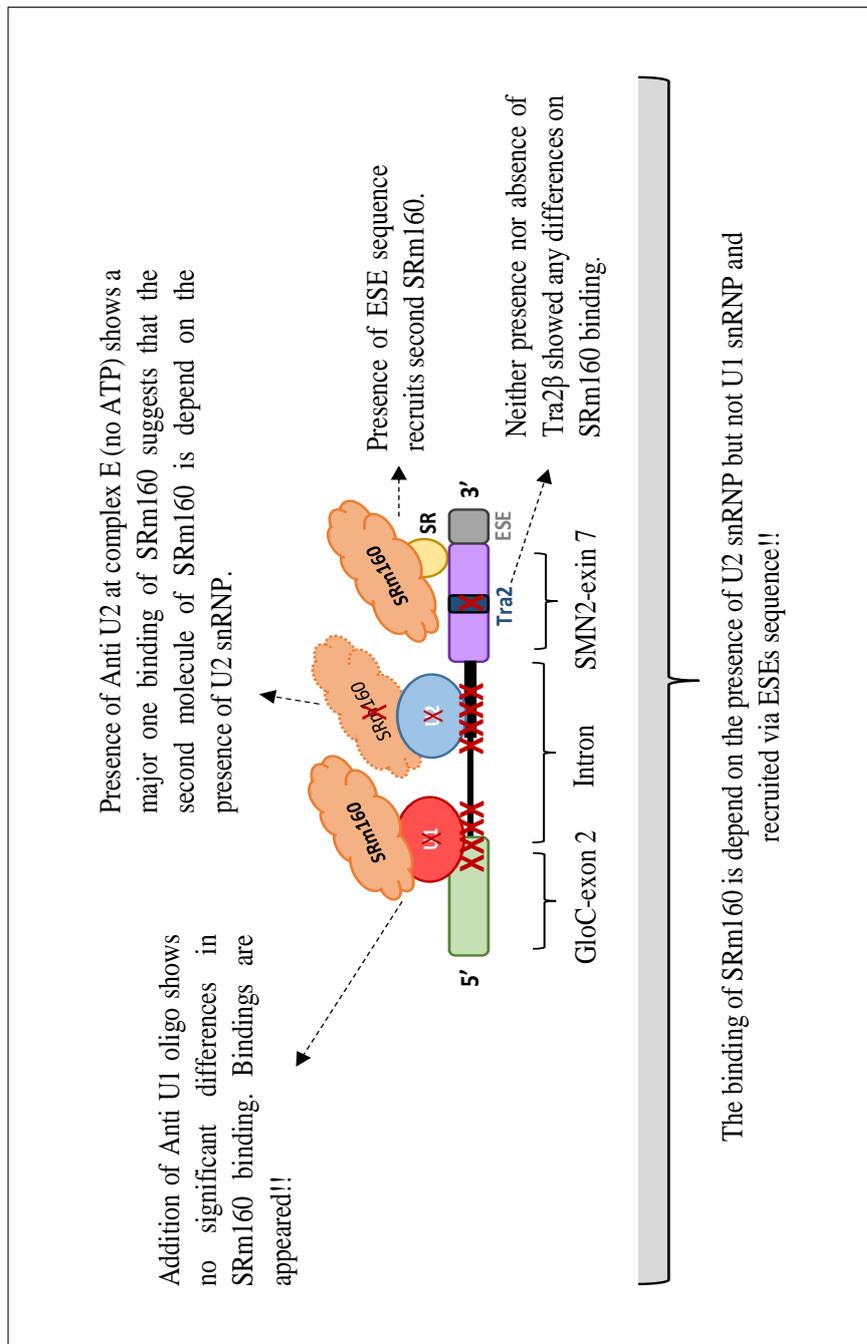


Figure 7.1: Summary diagram. - The state of SRm160 protein at single molecule level.

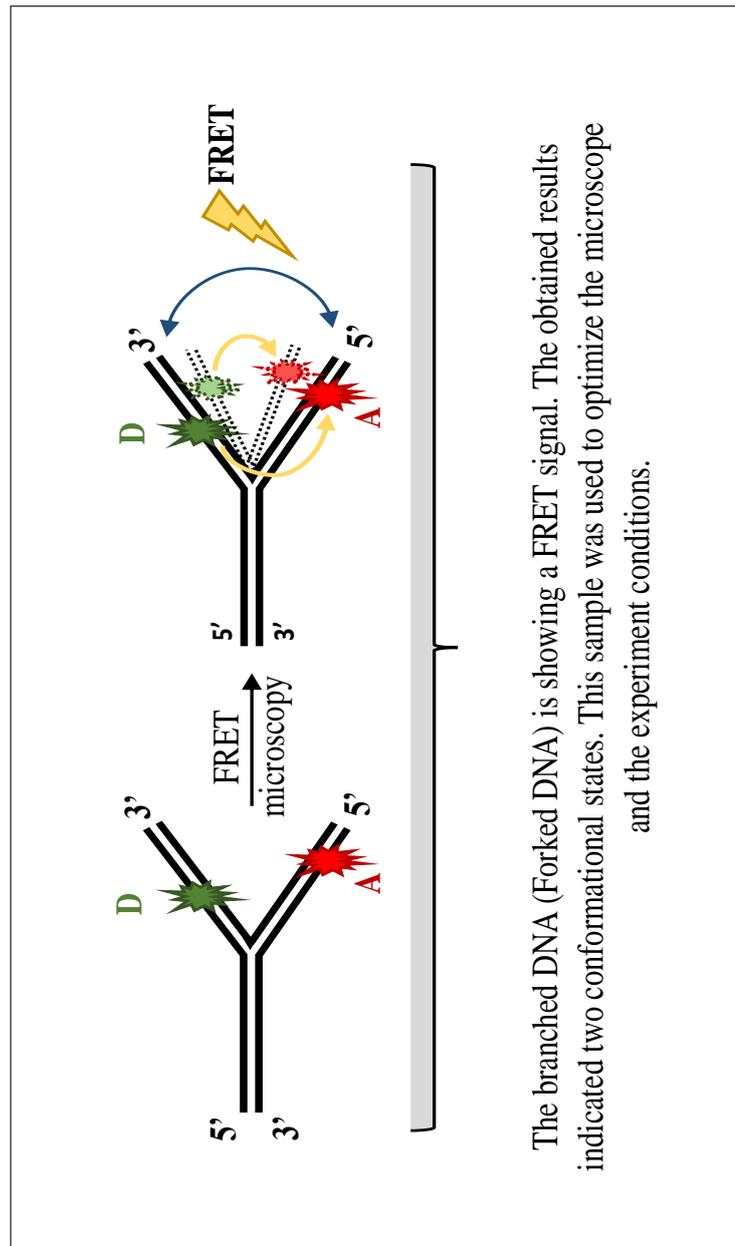


Figure 7.2: Summary diagram. - The state of Forked DNA at single molecule level through the fluorescence resonance energy transfer microscopy scope.

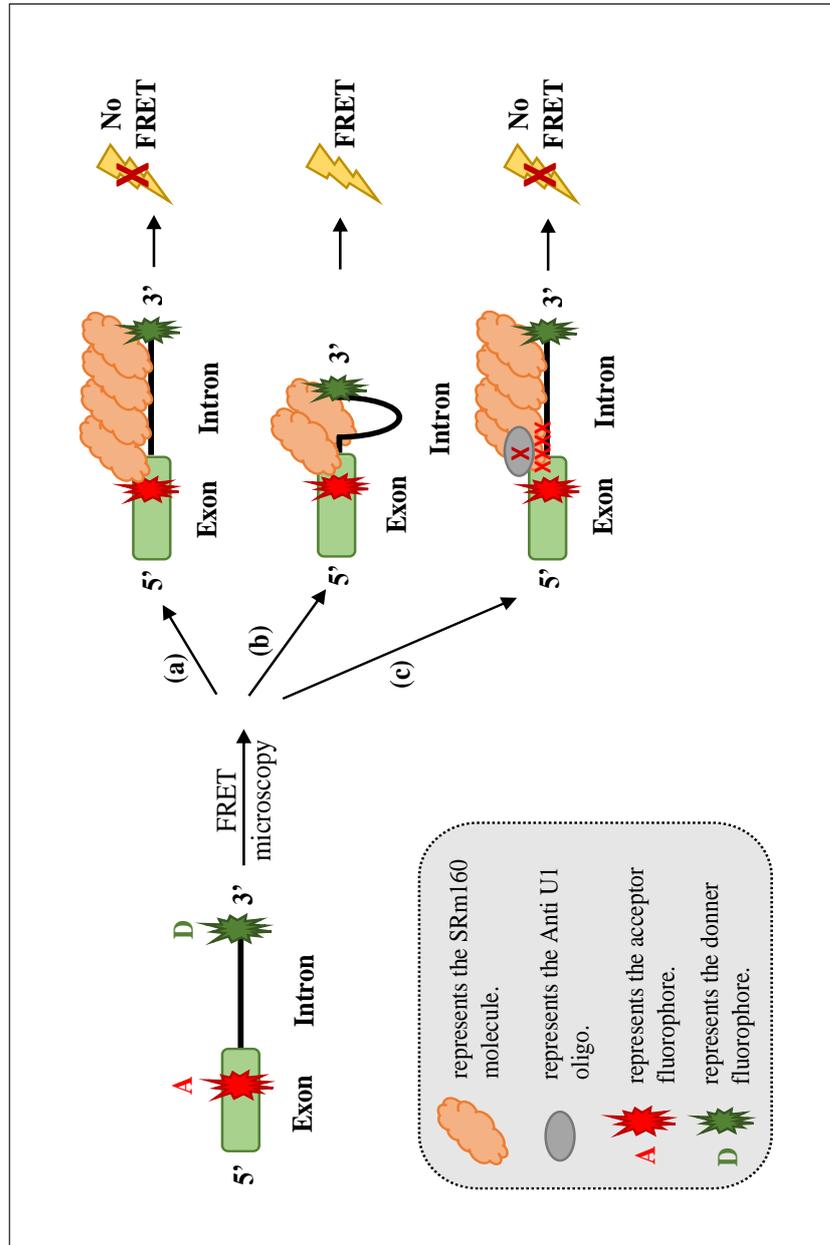


Figure 7.3: Summary diagram. - The state of Double-labelled RNA at single molecule level through the fluorescence resonance energy transfer microscopy in the presence of our home-made nuclear extract (SRm160).

### 7.1 The Microscope

Single molecule colocalisation measurements in combination with the analysis of intensity time traces is an essential process in order to investigate and identify the state of the binding of proteins of interest to RNA. It is also important obtaining quantitative information about the number of bound proteins.

Total internal reflection fluorescence (TIRF) is a perfectly suited technique for these measurement that allows a selective illumination of surface-bound molecules during a period of time. The TIRF microscope used in this research work was accomplished by Dr. Andrew Hudson from the Chemistry Department at The University of Leicester.

Before doing any experiments, the alignment is a fundamental step in order to use the TIRF microscope. In our home-made microscope, laser beams are generated separately and merged together into the beam combiner. The input and output of the beam are aligned to maximise the laser output to prevent the laser output to be weak.

After emerging from the laser combiner, the laser beams goes to the objective lens via a polarising beam splitter, a beam expander and a lens that focuses the beam on to the back focal plane of the objective. The objective then directs light to the surface of the slide chamber (cover slip). All of the mirrors involved were also aligned for accuracy detection. Furthermore, there are a few laser filters between the objective and the camera, depending on which laser is used.

Multiple beam splitters, filters and cameras were used for a simultaneous or sequential acquisition in order to separate any resulting fluorescent emission.

## 7.1 The Microscope

---

Once the laser beams reflect from the cover slip, they can excite fluorescent molecules immobilized on the cover slip. Those spots can be captured by the camera located on the right side or the bottom side of the microscope set up. There are two translational controls on the stage to adjust the position of the stage. This kind of control helps us to take the image from different positions of the slide.

Ethanol (about 80 %) and length tissues were used to wipe off any dirt on the laser filters or mirrors surfaces in order to ensure nothing accumulate on their surfaces and hence affect the recorded data (signals).

Settings on the LabView interface were also important. The (BM Gain) on EMCCD has to be set to 300 and has to be turned on to make sure there is enough light to be captured by the camera. In the (Binning) window, although the camera built into the microscope could take pictures of a  $512 \times 512$  pixel field, only  $256 \times 256$  pixel images were decided to be taken.

When preparing splicing samples for the TIRF microscope, it was important to adjust the intensity of the spots to an appropriate level. Also, it is important to adjust the number of spots on each acquisition to maximise the efficiency (see Chapter 3, section 3.5).

The presence of two different fluorescent dyes (fluorophores) in the sample of interest is the main key of using TIRF which allows us to employ a labelled protein (GFP-SRm160) with the labelled pre-mRNA (Cy5-pre-mRNA) for single molecule colocalisation experiments. It also allows us to employ the forked DNA (Double-labelled DNA) and the double-labelled RNA for FRET experiments.

The presence of fluorophores helps us to look at the relationship between the substrate (mRNA) and the protein of interest (SRm160) or to investigation

## 7.1 The Microscope

---

the state of fluctuations of a molecule of interest (forked DNA or double-labelled RNA).

Here in this work, regarding the colocalisation measurements, the chosen fluorophores were GFP (fluorescence in green) and Cy5 (fluorescence in red). Cy5 was easily attached to mRNA molecules as described previously and GFP was available in existing vector that allows us to clone them into our proteins of interest. The wavelength of 488 nm excites a single fluorophore of GFP while the wavelength of 561 nm excites Cy5. The chosen fluorophores for FRET measurements were Cy3 (fluorescence in green) and Cy5 (fluorescence in red) in the branched DNA. ATTO fluorophore (fluorescence in green) and Cy5 (fluorescence in red) were used in the double-labelled RNA.

Detection of the fluorophores was done sequentially in the colocalisation measurements by carefully choosing the order of excitation depending on the wavelengths used for excitation in order to avoid the excitation of a second fluorophore. An initial excitation with the longest wavelength excites only a single fluorophore. The illumination is continued until this fluorophore is bleached. A subsequent excitation with the next shorter wavelength for the next fluorophore is now applied. By this stage, the subsequent wavelength is restricted only to this second fluorophore. This sequence of excitation guarantees separate detection from the two employed fluorophores. While all FRET samples were illuminated via 532 nm (donor emission wavelength)

Moreover, the principle key in using the TIRF method is the intensity of the input signal (the intensity of the excitation wavelength) and hence the emission wavelengths. For this reason, the intensity of the signal has to be optimised. One way to optimise the signal intensity is to increase the intensity of the excitation which lead to an increased emission. Unfortunately, this step

## 7.1 The Microscope

---

shorts the time before the fluorophores bleach (fast bleaching) which might leads to intervals between individual bleaching steps that can not be resolved and unobserved bleaching during the initial frames and hence detection difficulty. Conversely, decreasing the intensity of the excitation results in a lower signal, complicating the detection of spots and hence the time traces analysis. Also, lowering the excitation intensity leads to a longer active time of the fluorophores before bleaching which increases the average time between bleaching events for an easier detection of steps in the intensity time traces. This leads to increasing the necessary acquisition times (each fluorophore has to be bleached completely).

Additionally, acquisition times should be enough to ensure a consistent state of the sample over the whole set of acquisitions and not to increase the amount of recorded data which makes data handling, storage and analysis more complicated.

Another side of optimising the signal intensity is the background noise. The ambient light in the room is a source of this noise as it is interference with the acquisition. This was avoided by covering the microscopic set-up with an opaque cover.

Optimising the length (period) of the recorded data is important too. Its related to the positional stability of the sample over time. This position has to be guaranteed in term of stability during the experiment. The implementation of an auto-focus compensated for the occurring changes. The ability to automatically focus in correspondence to the currently used wavelength provides the additional advantage of a quick and precise switch between different wavelengths without skipping the valuable first frames. Another important setting is the exposure time of the camera. The time exposure affects the

## 7.2 The Analysis

---

signal-to-noise ratio. Such that, longer exposure times lead to an increased signal-to-noise ratio. A good compromise was found with an exposure time of 0.1 s.

An important feature of the microscope is its interface with the computer. The lasers, the stage and the camera are all interfaced with the computer. The advantages of it lead to; increasing the ease of handling, allowing the implementation of the automatic focus, allows to automate the acquisition, also reducing the time needed to perform one set of experiments, minimising the potential for human errors during the acquisitional process and ensuring a high consistency between different acquisitions. Both connections with the computer and implementations were done using a software named LabView (national instruments). This program simplifies the interaction with the microscope, allows a quick familiarisation for new users and automatically creates documents of performed experiments for the later analysis of the data. This development was kindly accomplished by Dr. Robert Weinmeister.

In conclusion, optimizing the microscopic in terms of detection efficiency, the lowest possible background and noise levels is a fundamental process before performing any experiments. This optimisation ensures a reliable identification of colocalisation and bleaching steps. Overall, the performance of the set up with these various improvements resulted in a good quality of the acquisitions.

## 7.2 The Analysis

Following the optimisation of the microscope, performing the experiment, recording of all data and documentation of results, analysis of the obtained

## 7.2 The Analysis

---

results is the key to understand all findings. This analysis consists of several steps:

- a. images have to be assigned to a wavelength,
- b. spots have to be identified,
- c. their positions have to be determined,
- d. colocalised spots have to be assigned to each other,
- e. the intensity time traces of the colocalized spots have to be generated,
- f. the number of bleaching steps have to be analysed,
- g. the resulting formation has to be documented and presented in an easy way,
- h. documented results have to be presented in a clear and easy way to understand.

Regarding analysis, employing algorithms for these steps opens the door for further automation.

In the end, the output from the all algorithms used had to be refined to enhance the quality of the step assignment. The implementation of the data analysis with these algorithms was done in Matlab. Algorithms and MatLab were modified by Dr. Robert Weinmeister.

Next, the presentation of the data is done as a bar chart in order to quickly convey the information extracted from the recorded data. The percentage of SRm160 proteins bind to pre-mRNA is shown on the X axis. The distribution of the number of fluorescent proteins present can vary at each

### 7.3 The Preparation

---

distinct complex (Y axis). The number of marker spots and the number of colocalised spots is shown to give an indication of the statistical relevance of the results shown. The category for uncertain assignments of bleaching steps, represented as X, was included not to exclude colocalization events where the intensity time traces are of poor quality.

### 7.3 The Preparation

It was decided to prepare nuclear extracts from 293T cells with GFP-SRm160. The expression levels were performed via western blot experiment in compare to an available commercial nuclear extract without a labelled protein. To prepare nuclear extracts containing fluorescent-labelled SRm160, the mEGFP-SRm160 plasmid was transfected into 293T cells. The transfected cells were then used to make nuclear extracts. The labelling efficiency for GFP-SRm160 nuclear extract was high enough ensuring that the SRm160 proteins are almost labelled via GFP.

Optimising the annealing conditions of pre-mRNA substrates was the next step. Conditions used for this purposes allow for almost no free Cy5-oligo present as seen on a denaturing gel. This was comparable with the colocalisation experiments. Analysis of an experiment that allows the formation of the early complex E or complex A via calculating the un-colocalised spots (Cy5-pre-mRNA) revealed almost one main binding which indicates that each single molecule of RNA is labelled via Cy5-oligo. Also, SRm160 was found to not bind Cy5 oligo in the absence of pre-mRNA. These results guarantee that the presence of any free-Cy5 oligo somehow does not affect the binding of SRm160 to pre-mRNA. Also, any remaining amount of free pre-mRNA with-

## 7.4 The Protein; SRm160

---

out any fluorescent label does not affect the results as it is unlabelled and not visible and does not interfere with the splicing progress of labelled pre-mRNA.

Labelled pre-mRNA showed its ability to immobilize to the surface of the un-modified slide in the presence of a nuclear extract. This attachment does not increase in the presence of a streptavidin. As the presence of a streptavidin can provide binding sites for the biotin end of Cy5 oligo, it was logically expected that the presence of streptavidin on top of a slide chamber increases the amount of attached pre-mRNA confined to the surface. However, this was not found to be the case. Two explanation might exist;

- a. The biotin of the end of the Cy5 oligo is not accessible to the streptavidin
- b. the activity of the biotin at the end of the Cy5 oligo is inactivated by the protein complex forming around the pre-mRNA.

## 7.4 The Protein; SRm160

### 7.4.1 The SM-Colocalisation Vs. SRm160

The binding sites of SRm160 on pre-mRNA, its role and presence during the splicing reaction are so far unknown. There is nothing known about the actual number of bound SRm160 during splicing at a distinct splicing complex. Here in this research work, we identified the behaviour of how many SRm160 proteins are present, the possible factors that influence their numbers to gain a deeper understanding and a clearer investigation of the role of SRm160 protein.

As the complex formation of individual pre-mRNA molecules can occur at different rates during splicing, we stalled the formation of splicing at the early complex E and complex A following our group procedures. All conditions used in this research allowed an efficient stalling of the complex formation.

## 7.4 The Protein; SRm160

---

Stalling the splicing progress at the early complex E was performed by a depletion of ATP. This step led to proportions of one, two, or a multitude of SRm160 molecules present (12% of the Glo C pre-mRNA were found to contain multiple bound molecules of SRm160). The prominent occurrence of a multiple numbers of the SRm160 present could be referred to as aggregates. These aggregates are not seen in the absence of pre-mRNA, even after RNase treatment.

The addition of ATP and anti-U6 oligo stalled the splicing at complex A. The pattern is dramatically different. Strikingly, all complexes contain one to two molecules of SRm160 accompanied by increasing in colocalisation. This leads to kind of breakup of aggregates presented at the early complex E.

The amount of bound SRm160 in the early complex E and complex A stands out. Most research papers showed that SRm160 protein exists only in the lariat product and the 5'-exon product. Additionally, there is no research paper identified that SRm160 protein can bind pre-mRNA at the early complex E at least. This alludes to an existing mechanism behind our findings (the binding of SRm160 at a splicing complexes).

The addition of a nucleotide complementary to U1 snRNP at condition supporting the formation of the early complex E demonstrate that the distribution binding of SRm160 does not differ from the one obtained in the absence of this oligo (anti-U1 snRNP). No significant differences exist even in the percentage of the colocalisation (dropped from 12% at the early complex E to 10% at same condition but with the addition of anti-U1 oligo). This result that the binding of SRm160 does not depend on the presence of U1 snRNP as SRm160 still binds pre-mRNA in condition blocking the binding of U1 snRNP. This result argues against the idea that SRm160 proteins are highly dependent

## 7.4 The Protein; SRm160

---

on the presence of U1 snRNP and stabilised via the presence of U2 snRNP. It was and still believed that U1 snRNP mediates the interaction between SRm160 molecules and the pre-mRNA as SRm160 does not consist of RNA binding domains (RRM).

The U1 snRNP binds to the 5' splice site and selects the 5' splice site in conjunction with SRm160. A random amount of SRm160 is bound to the pre-mRNA without a strong consensus 5' splice site present in GloM and simultaneously blocking of the U1 snRNP with a large proportion of aggregates. The binding distribution of SRm160 in conditions that support the formation of the early complex E has no clear definitions as is dominated by accumulations of SRm160 at the pre-mRNA. In the presence of ATP and anti-U6 oligo, there is a significant increase in the binding to one bound SRm160. This condition lowers the amount of aggregates and increases the colocalisation efficiency (from 24% to 33%). The aggregates still form at levels equal to Globin C in the early complex E and complex A without the blocked U1 snRNP. The absence of any significant changes regarding the binding of SRm160 in both complexes with blocked U1 snRNP nor with the addition of anti-U1 snRNP or by using GloM substrate seemingly independent of the presence of a strong consensus 5' splice site and hence the presence of U1 snRNP.

Although GloM construct was supposed not to form complex A because it lacks a strong consensus 5' splice site, incubation with ATP does clearly lead to multiple bound SRm160 proteins. This ensures that SRm160 proteins not only depends on the presence of U1 snRNP but there is something else that affects binding or recruiting more molecules of SRm160 proteins.

The RS domain of SRm160 is heavily phosphorylated and phosphorylation changes of SRm160 play an important role in splicing. Incubation with

## 7.4 The Protein; SRm160

---

PhosStop at the early complex E led to no changes in the binding distribution of SRm160. According to this result, the appearance of no significant changes of SRm160 binding introduced by the addition of ATP might be linked to phosphorylation changes.

The fact that PhosStop did not show any effect on SRm160 itself suggests that the phosphorylation of SRm160 itself does not play a role at this stage. The phosphorylation of SRm160 might be involved in other controlling mechanisms.

SRm160 in the nuclear extract used is overexpressed compared to untransfected nuclear extracts. We found that the overexpression of SRm160 shifts the binding pattern towards higher colocalisation with a distribution dominated by aggregates. Such that, the aggregates might be thought due to the overexpression of SRm160. By using an abundance of available commercial nuclear extracts in order to decrease the ratio of SRm160 showed no significantly different binding pattern in both the early complex E and complex A. The only difference was the smoothed pattern. We conclude from these results that our levels of expression of SRm160 did not affect the observed binding patterns and hence our findings.

For the construct SMN2, we investigated whether or not the presence of the ESE increased the splicing efficiency and the number of SRm160 molecules bound to the pre-mRNA. The number of bound SRm160 is generally seen to increase as a splicing enhancer is increases. This relation was demonstrated to be influenced by specific sequences, such as GAA-repeats sites. We found that in the absence of ATP, pre-mRNA associates with a multiple numbers of GFP-SRm160 molecules. The number of bound GFP-SRm160 is strikingly increased with increasing in colocalisation percentage (from 13% to 24%) in

## 7.4 The Protein; SRm160

---

SMN2 containing an ESE sequence. In the presence of ATP, binding of GFP-SRm160 is restricted to two defined molecules in the presence of an ESE and one to two molecules in the absence of an ESE. These results indicate that the number of GFP-SRm160 that bind pre-mRNA substrate is highly dependent on the ESE sequence. Additionally, SMN2 with and without ESE sequences were used in extract pre-incubated reaction to deplete ATP in the presence of anti-U1 oligo. This experiment showed that the association of GFP-SRm160 with SMN2 substrate does not depend on U1 snRNP even in the presence of an ESE sequence. These findings conflict with the previous observation that the SRm160 does not bind to pre-mRNA in the absence of U1 snRNP. Unexpectedly, a single GFP-SRm160 molecule was identified when anti-U2 oligo is present. These results demonstrate various points; the association of SRm160 with SMN2 pre-mRNA is enhanced via ESE sequences, and is highly dependent on U2 snRNP but not U1 snRNP.

Other factor might affect the recruitment of SRm160 molecules is the presence of Tra2 $\beta$ . SMN2 substrate without both Tra2 $\beta$  and ESE sequences is used in comparison to SMN2 without Tra2 $\beta$  site containing ESE sequence. We found that after the incubation in nuclear extract pre-incubated to deplete ATP, 11% and 14% of the RNA complexes contained multiple molecules of GFP-SRm160 were identified in the absence and the presence of ESE sequence, respectively. Phostop treatment in the absence of ATP resulted in single defined GFP-SRm160 molecule and a single molecule in the absence of the ESE sequence, and one to two molecules in the presence of ESE sequence. With ATP, the presence of ESE sequence recruited more GFP-SRm160 molecules. Two GFP-SRm160 molecules were identified. The colocalization efficiency was increased from 9% in the absence of ESE to 19% in the presence of the ESE

## 7.4 The Protein; SRm160

---

sequence. These results indicate again that the presence of ESE sequences recruit more GFP-SRm160 in SMN2 substrate which conflicts with the previous studies that the association of SRm160 protein is specific for ESE-dependent splicing containing GAA repeats.

Moreover, the addition of anti-U1 oligo to ATP-depleted reaction, the binding of SRm160 to SMN containing ESE sequences is slightly higher with similar distribution pattern in comparison to SMN2 in the absence of this oligonucleotide. These results indicate the absence of U1 snRNP does not affect the binding of GFP-SRm160 to pre-mRNA. Whereas in the absence of ESE sequence, one to two GFP-SRm160 molecules were identified. This supports the idea that level of GFP-SRm160 is enhanced by the presence of ESE sequences even in the presence of an oligonucleotide complementary to U1 snRNA. The presence of an anti-U2 oligo significantly reduced the occurrence of single molecule in the great majority of complexes formed in the absence of ATP. This result in conjugation with the low colocalization efficiency suggests that the binding of SRm160 is associated with U2 snRNP. In summary, the number of GFP-SRm160 is enhanced via ESE sequences. The absence of U1 snRNP had no affect on the binding distribution. Integrally, the association of GFP-SRm160 to pre-mRNA is not stimulated by the Tra2 $\beta$  site. All these findings conflict with the previous studies which demonstrated that that hTra2 $\beta$  of *Drosophila* Tra2 $\beta$  bind to ESEs containing GAA repeats and, in conjugation with SR family proteins such as SRm160, promote ESE-dependent splicing.

As from all previous studies, the association of SRm160 to pre-mRNA is dependent on the presence of U2 snRNP. If this is the case, the absence of splicing signals might affect this interaction. SRm160 is supposed not to

## 7.4 The Protein; SRm160

---

bind to spliced RNA (intron-less substrate) as there are no intron sequences and hence no more U2 snRNP exists. Intron-less constructs were used to investigate whether SRm160 is following this logic or not. Interestingly, single molecule colocalisation experiments showed unexpected results in which SRm160 proteins are still bound to the intron-less GloC construct in the early complex E and complex A condition. In the absence of ATP, the distribution pattern of SRm160 bound to the spliced RNA was very similar to GloC with a high percentage of colocalisation efficiency. Similar results to GloC were also observed in the addition of ATP. However, the presence of binding of SRm160 in the absence of U2 snRNP; where the intron sequences is deleted, the occurrence of the binding of SRm160 molecules might be due to the presence of ESEs sites in the exons which is consistent with the results that the binding of SRm160 is influenced by the presence of ESEs sites. SRm160 needs its RS domain to participate in splicing complexes formation. This is in agreement with a published paper which indicates that the SRm160 could bind RNA via a region named PWI. It was found that the U1 snRNP is required for the interaction of SRm160 to pre-mRNA, but our findings suggest that most likely this is not the case because SRm160 proteins are bound to pre-mRNA in a condition that allows the blocking of U1 snRNP. We also found a clear effect of the ESE sequences on recruiting more molecules of SRm160 proteins. We did see a clear correlation between the binding of SRm160 and the presence of ESE in SMN2 substrate. This finding is in-agreement with most published papers which indicate that this protein is effectively reacted via specific sequences of ESE sites in specific substrates.

## 7.4 The Protein; SRm160

---

### 7.4.2 The FRET signal Vs. SRm160

FRET technique is a gate for a deeper investigation of a molecule of interest in a real time. This technique allows hidden information in the ensemble methods to be investigated.

The microscope set up was optimised as mentioned previously in this chapter. Slides (cover slips) were also treated and modified to ensure an efficient immobilisation of the sample of interest. The impurities of both slides and reagents were also checked. The cross-talk signals were corrected. All experiment conditions (concentrations and dilutions) were optimised too. All those processes were performed using a branched DNA (forked DNA). As this substrate has two suitable fluorescent dyes, FRET technique was implemented. The obtained results indicate FRET signals. The FRET signal appeared in two states. Those two states of FRET are a point of a molecule fluctuations. Those results were obtained via both the prism-based TIRF microscope and the objective-based TIRF microscope. Overall, the Holliday junction substrate (Forked DNA) was a good model for use in FRET experiments as a control.

As SRm160 is the main feature of this research work, more investigations were initiated in parallel to the single molecule colocalisation measurements. The state of SRm160 was followed at single molecule level but with the FRET technique using a suitable double-labelled substrate.

A double-labelled RNA was used. Experiments were performed under different conditions. The obtained results indicate no FRET signal was recorded at a condition supporting the formation of the early complex E (no ATP). The absence of the FRET signals might refer to two options: (1) presence of multiple bound-factors to the substrate which might restrict the

## 7.5 Summary

---

fluctuation and flexible movement and/or (2) the presence of a number of a huge molecules of SRm160 which could affect the energy transfer between the two fluorophores via limiting the movement of the substrate. This finding was comparable to the single molecule colocalisation results. In the absence of ATP, multiple molecules of SRm160 were bound to the substrate. On the other hand, FRET signal was appeared at the condition supported the formation of complex A (the addition of ATP and anti-U6 oligo). At this conditions, the binding of SRm160 was controlled. This binding was identified to be decreased in comparison to the early complex E condition through the colocalisation measurements. Those findings together indicate that the SRm160 molecule affected the appearance of fluctuations by and hence affecting the FRET signal. As the number of SRm160 molecules is decreased, the greater opportunity for the FRET signal to be detected. Additionally, the presence of a high signal of FRET confirms the de-association step of some bound-factors presented at the early complex E. Interestingly, the presence of  $\alpha$ U1 (no U1 snRNP should occur at this stage) oligo affected the FRET signal. The obtained results showed almost no FRET signal was detected. This could be due to the remaining bound-molecules of SRm160 to the substrate which was also investigated via the single molecule colocalisation experiments. Overall, the 160 kDa molecule (SRm160) and the state of its binding to a substrate affected the FRET signals and hence the substrate fluctuations.

## 7.5 Summary

The purpose of this research was to investigate the state of SRm160 protein during the splicing reaction under conditions supported the formation of the

## 7.5 Summary

---

early complex E, or complex A and how its binding influenced via the presence of other factors.

SRm160 was studied since 1994. It was found that SRm160 is a splicing co-activator factor. It is localised in peackles. It was believed that SRm160 protein acts in conjugation with SR300 and it was found that SRm160 is much more important than SR300. SRm160 is the almost the leader. It was also believed that SRm160 protein is recruited via ESEs sequences sites. It was demonstrated that the recruitment of SRm160 protein was increased as the number of those ESEs sequences are increased. It was also demonstrated that the relationship between the recruitment and the number of ESEs is related to a specific sequences (GAA-repeats). It was also found that SRm160 molecules are recruited via Tra2 $\beta$  sequences sites.

However, All previous experiments were performed via an assembled methods via a specific antibody. All published papers concluded that SRm160 reacts as one rigid molecule binding the 5' ss via U1 snRNP, the 3' ss via U2 snRNP and the ESE sequence via SR proteins. It was identified that the SRm160 binding depends on U1 snRNP and stabilizes via U2 snRNP. No single molecule experiments were done so far. From this point, the investigation of the state of SRm160 is completed.

The state of SRm160 protein is as the following:

- a. Single molecule colocalisation experiments was performed here in this work to look at the state of binding of SRm160 protein.
- b. The binding of SRm160 protein to GloC construct showed a multiple binding at the early complex E (no ATP) and complex A (with ATP and with  $\alpha$ U6 oligo).

## 7.5 Summary

---

- c. The binding of SRm160 protein to GloC construct showed no significant differences in the binding distribution when the anti-U1 oligo is added.
- d. The previous point was confirmed by using GloM construct (the 5' ss is blocked), the obtained results showed no significant differences in binding distribution.
- e. We concluded that SRm160 molecules remain bind to the RNA construct in the presence and the absence of U1 snRNP. This finding argues with the previous investigations (in the literature).
- f. The binding of SRm160 protein to the intron-less substrate showed no significant differences from the normal GloC.
- g. The binding of SRm160 protein to SMN2 construct showed an increasing level when the ESE site is involved (multiple conditions were tested, refer to Chapter 5).
- h. The binding of SRm160 protein to SMN2 construct showed no significant differences when anti-U1 is added to the reaction.
- i. The binding of SRm160 protein to SMN2 construct showed one single bleaching step (a single bound SRm160) when anti-U2 is added to the reaction.
- j. The binding of SRm160 protein to SMN2 construct showed that the association of SRm160 to pre-mRNA does not stimulated by Tra2 $\beta$ . Overall, the presence and the absence of Tra2 $\beta$  site does not significantly affected the state of binding of SRm160.

## 7.5 Summary

---

- k. We concluded that SRm160 molecules showed a significant differences regarding the binding to RNA construct only in two cases; the absence of U2 snRNP and the presence of ESE sequences sites. This finding is totally argues with all what was believed on regarding the biding of SRm 160 proteins.

Moreover, it was very interesting to investigate whether a rigid molecule such as SRm160 can have an influence on the FRET signals at a distinct splicing complex. Results performed in this work concludes the following:

- a. No FRET signal was appeared at the condition supporting the formation of the early complex E (no ATP). This is comparable with the colocalisation results which indicates a multiple binding of SRm160 molecules alongside the pre-mRNA substrate. This means that the presence of a multiple number of a huge molecule (SRm160) prevent the energy transfer. Additionally, this result could refer to the limitation of the possibility of fluctuation which caused by the presence of this rigid molecule.
- b. FRET signal was appeared at the condition supporting the formation of complex A (+ATP+ $\alpha$ U6 oligo). Also, this finding was in agreement the colocalisation results at the same condition. The phosphorylation was decreased the number of SRm160 proteins. This means the energy transfer between the two flurophores is nearly possible.
- c. The presence of anti-U1 oligo prevent the FRET signal to be detected. This means that SRm160 were remain bind to the construct. Remaining of a huge molecules affected the FRET signal as identified at the early complex E condition. This findings confirms our conclusion that the binding of SRm160 does not depend on the presence of U1 snRNP.

## 7.6 Future Work

As all constructs used in this work consist of a strong 5' exon and one intron, it will be interesting to investigate whether the binding behavior of SRm160 protein depends on the presence of those ESEs in the 5' exon or not. This hypothesis is established in this thesis. A series of GloC substrates were used. These substrates differed in the length of the 5' exons. They were designed from GloC. Standard splicing reaction was implemented. Overall, short exons substrates used here in this work shows a splicing signal with different efficiencies whether ESEs exist or not. These findings raised a thought that the ESE sites at the 5' exon might be a reason for bound-SRm160 molecules remaining on a substrate that lacks a 3' ESE or the intron or the 5' ss. By this stage, using a very short substrate that lacks ESE sites in both the 5' and 3' exons, e.g., a substrate with only two nucleotides in the two locations (5' and 3' exons) might outline another investigation regarding the effect of the presence of ESE site to the binding state of SRm160.

Moreover, the enhancement of SRm160 binding which was observed in the presence of ESE sites raised a question; does the number of ESE sites affect the binding? Or does the increasing number of ESEs can recruit more molecules of SRm160? To investigate this hypothesis; SMN2 substrates with some ESEs might be the goal in the future. Additionally, Sequencing the RNA through the genome-wide will be a good step to find where SRm160 is bound the RNA and which type of RNA are suitable to recruit SRm160 molecules. Also, analyzing the interaction of SRm160 with RNA substrates to precisely locate where exactly SRm160 is bound the RNA by using cross-linking immunoprecipitation (CLIP) technique is a significant forward step in

## 7.6 Future Work

---

the future.

These experiments with what I determined earlier in this thesis, might outline a molecular framework for the binding of the SRm160 protein to pre-mRNA in splicing reaction.

# Chapter 8

## Appendices

## 8.1 Sequences of RNA Constructs

---

### 8.1 Sequences of RNA Constructs

#### 8.1.1 Glo C

[H]

---

1 **GGGCTGCTGG** **TTGTCTACCC** **ATGGACCCAG** **AGGTTCTTCG** **AGTCCTTTGG**  
51 **GGACCTGTCC** **TCTGCAAATG** **CTGTTATGAA** **CAATCCTAAG** **GTGAAGGCTC**  
101 **ATGGCAAGAA** **GGTGCTGGCT** **GCCTTCAGTG** **AGGGTCTGAG** **TCACCTGGAC**  
151 **AACCTCAAAG** **GCACCTTTGC** **TAAGCTGAGT** **GAAGTGCAGT** **GTGACAAGCT**  
201 **GCACGTGGAT** **CCTGAGAACT** **TCAGGTAAGT** **TTGGGGACCC** **TTGATTGTTC**  
251 **TTTCTTTTTT** **GCTATTGTAA** **AAATTCATGT** **TATATGGTCG** **ACTCTGCTAA**  
301 **CCATGTTTAT** **GCCTTCTTCT** **TTTTCCTACA** **GCTCCTGGGC** **AACGTGCTGG**  
351 **TTATTGTGCT** **GTCTCATCAT** **TTTG**

#### 8.1.2 Globin M

---

1 **GGGCTGCTGG** **TTGTCTACCC** **ATGGACCCAG** **AGGTTCTTCG** **AGTCCTTTGG**  
51 **GGACCTGTCC** **TCTGCAAATG** **CTGTTATGAA** **CAATCCTAAG** **GTGAAGGCTC**  
101 **ATGGCAAGAA** **GGTGCTGGCT** **GCCTTCAGTG** **AGGGTCTGAG** **TCACCTGGAC**  
151 **AACCTCAAAG** **GCACCTTTGC** **TAAGCTGAGT** **GAAGTGCAGT** **GTGACAAGCT**  
201 **GCACGTGGAT** **CCTGAGAACT** **TCAGGTTTGG** **GGACCCTTGA** **TTGTTCTTTT**  
251 **TTTTTTCGCTA** **TTGTAAAAAT** **TCATGTTATA** **TGGTCGACTC** **TGCTAACCAT**  
301 **GTTTCATGCCT** **TCTTCTTTTT** **CCTACAGCTC** **CTGGGCAACG** **TGCTGGTTAT**  
351 **TGTGCTGTCT** **CATCATTTTG** **G**

## 8.1 Sequences of RNA Constructs

---

### 8.1.3 Globin C GG

---

1 **GGGCTGCTGG** **TTGTCTACCC** **ATGGACCCAG** **AGGTTCTTCG** **AGTCCTTTGG**  
51 **GGACCTGTCC** **TCTGCAAATG** **CTGTTATGAA** **CAATCCTAAG** **GTGAAGGCTC**  
101 **ATGGCAAGAA** **GGTGCTGGCT** **GCCTTCAGTG** **AGGGTCTGAG** **TCACCTGGAC**  
151 **AACCTCAAAG** **GCACCTTTGC** **TAAGCTGAGT** **GAAGTGCAGT** **GTGACAAGCT**  
201 **GCACGTGGAT** **CCTGAGAACT** **TCAGGTAAGT** **TTGGGGACCC** **TTGATTGTTC**  
251 **TTTCTTTTTT** **GCTATTGTAA** **AAATTCATGT** **TATATGGTCG** **ACTCTGCTAA**  
301 **CCATGTTTAT** **GCCTTCTTCT** **TTTTCCTACG** **GCTCCTGGGC** **AACGTGCTGG**  
351 **TTATTGTGCT** **GTCTCATCAT** **TTTGG**

### 8.1.4 Rabbit $\beta$ -globin Sequence

---

1 **GGGCTGCTGG** **TTGTCTACCC** **ATGGACCCAG** **AGGTTCTTCG** **AGTCCTTTGG**  
51 **GGACCTGTCC** **TCTGCAAATG** **CTGTTATGAA** **CAATCCTAAG** **GTGAAGGCTC**  
101 **ATGGCAAGAA** **GGTGCTGGCT** **GCCTTCAGTG** **AGGGTCTGAG** **TCACCTGGAC**  
151 **AACCTCAAAG** **GCACCTTTGC** **TAAGCTGAGT** **GAAGTGCAGT** **GTGACAAGCT**  
201 **GCACGTGGAT** **CCTGAGAACT** **TCAGGTTGAG** **TTTGGGGACC** **CTTGATTGTT**  
251 **CTTCTTTTTT** **CGCTATTGTA** **AAATTCATGT** **TATATGGTCG** **ACTCTGCTAA**  
301 **CCATGTTTAT** **GCCTTCTTCT** **TTTTCCTACA** **GCTCCTGGGC** **AACGTGCTGG**  
351 **TTATTGTGCT** **GTCTCATCAT** **TTTGGCAGGT** **AAGTT**

## 8.1 Sequences of RNA Constructs

---

### 8.1.5 CEC

---

1 GGGCTGCTGG TTGTCTAGGG CTGCTGGTTG TCTACCCATG GACCCAGAGG  
51 TTCTTCGAGT CCTTTGGGGA CCTGTCCTCT GCAAATGCAG TTATGAACAA  
101 TCCTAAGGTG AAGGCTCATG GCAAGAAGGT GCTGGCTGCC TTCAGTGAGG  
151 GTCTGAGTCA CCTGGACAAC CTCAAAGGCA CCTTTGCTAA GCTGAGTGAA  
201 CTGCACTGTG ACAAGCTGCA CGTGGATCCT GAGAACTTCA GGTAAGTTTG  
251 GGGACCCTTG ATTGTTCTTT CTTTTTCGCT ATTGTAAAAA TTCATGTTAT  
301 ATGGTCGACT CTGCTAACCA TGTCATGCC TTCTTCTTTT TCCTACAGCT  
351 CCTGGGCAAC CTGCTGGTTA GGCCAGAGG TTCTTCGAGT CCTTTGGGGA  
401 CCTGTCCTCT GCAAATGCTG TTATGAACAA TCCTAAGGTG AAGGCTCATG  
451 GCAAGAAGGT GCTGGCTGCC TTCAGTGAGG GTCTGAGTCA CCTGGACAAC  
501 CTCAAAGGCA CCTTTGCTAA GCTGAGTGAA CTGCACTGTG ACAAGCTGCA  
551 CGTGGATCCT GAGAACTTCA GGTAAGTTTG GGGACCCTTG ATTGTTCTTT  
601 CTTTTTCGCT ATTGTAAAAA TTCATGTTAT ATGGTCGACT CTGCTAACCA  
651 TGTCATGCC TTCTTCTTTT TCCTACAGCT CCTGGGCAAC CTGCTGGTTA  
701 TTGTGC

### 8.1.6 Primer 16

---

1 AAATTAATAC GACTCACTAT AGGG-CTGCT GGTTGTCTAC CCA

### 8.1.7 Primer 17

---

1 AACTTAC-CT GCCAAAATGA TGAGACAGCA TGAGACAGCA

## 8.1 Sequences of RNA Constructs

---

### 8.1.8 SMN2

---

1 **GGGCTGCTGG** **TTGTCTACCC** **ATGGACCCAG** **AGGTTCTTCG** **AGTCCTTTGG**  
51 **GGACCTGTCC** **TCTGCAAATG** **CTGTTATGAA** **CAATCCTAAG** **GTGAAGGCTC**  
101 **ATGGCAAGAA** **GGTGCTGGCT** **GCCTTCAGTG** **AGGGTCTGAG** **TCACCTGGAC**  
151 **AACCTCAAAG** **GCACCTTTGC** **TAAGCTGAGT** **GAACTGCACT** **GTGACAAGCT**  
201 **GCACGTGGAT** **CCTGAGAACT** **TCAGG**GTGAG TTTGGGGACC CTTGATTGTT  
251 CTTTCTTTTT CGCTATTGTA AAATTCATGT TATATG**GTCG** **AC**AGACTATC  
301 AACTTAATTT CTGATCATAT TTTGTTGAAT AAAATAAGTA AAATGTCTTG  
351 TGAAACAAAA TGCTTTTTAA CATCCATATA AAGCTATCTA TATATAGCTA  
401 TCTATATCTA TATAGCTATT TTTTTTAACT TCCTTTATTT TCCTTACAG**G**  
451 **GTTT****TAGACA** **AAATCAAAAA** **GAAGGA**AGGT **GCTCACATTC** **CTTAAAT****CAG**  
501 **GAG**TAAGTCT GCCAGCATTA TGAAAGTGAA TCTTACTTTT GTAAAACTTT  
551 ATGGTTTGTG GAAAACAAAT GTTTTTGAAC ATTTAAAAAG TTCAGATGTT  
601 AGAAAGTTGA AAGGTTAATG TAAAACAATC AATATTAAG AATTTTGATG  
651 CCAA~~AA~~CTAT TAGATAAAAAG GTTAATCTAC ATCCCTACTA GAATTCTCAT  
701 ACTTAACTGG TTGGTTGTGT GGAAGAAACA TACTTTTACA ATAAAGAGCT  
751 TTAGGATATG ATGCCATTTT ATATCAC**gtc** **gac**TCTGCTA ACCATGTTC  
801 TGCCTTCTTC TTTTTCCTAC AG**CTCCTGGG** **CAACGTGCTG** **GTTATTGTGC**  
851 **TGTCTCATCA** **TTTTGGCAGG** **TAAGTT**

### 8.1.9 Reverse Tra2 $\beta$ deletion (SMN2)

---

1 ATTTAAGGAA TGTGAGCACC **GATTTTGTCT** **AAAACCCTGT** **AAGGAAAAAA**  
51 **GGAAAA**

## 8.1 Sequences of RNA Constructs

---

### 8.1.10 Reverse +ESE

---

1 **CTTCCTCCGC** **CTTGCTTCCT** **CCGCCTTG**AT TTAAGGAATG TGAGCACCTT  
51 CCTTCTTTTT GA

### 8.1.11 Reverse Tra2 $\beta$ +ESE (SMN2)

---

1 **CTTCCTCCGC** **CTTGCTTCCT** **CCGCCTTG**AT TTAAGGAATG TGAGCACCG**A**  
51 **TTTTGTCTAA** **AACCCTGTAA** **GGAAAA**

### 8.1.12 Frw.primers no (1)

---

1 **AAATTAATAC** **GACTCACTAT** **AGGGACCCAG** **AGGTTCTTCG** **A**

### 8.1.13 Frw.primers no (2)

---

1 **AAATTAATAC** **GACTCACTAT** **AGGGTCCTTT** **GGGGACCTGT** **CCT**

### 8.1.14 Frw.primers no (3)

---

1 **AAATTAATAC** **GACTCACTAT** **AGGGCTGCAA** **ATGCTGTTAT** **GAACAAT**

## 8.1 Sequences of RNA Constructs

---

### 8.1.15 Frw.primers no (4)

---

1 AAATTAATAC GACTCACTAT AGGGAATCCT AAGGTGAAGG CTCA

### 8.1.16 Frw.primers no (5)

---

1 AAATTAATAC GACTCACTAT AGGCAAGAA GGTGCTGGCT GGTGCTGGCT

### 8.1.17 Frw.primers no (6)

---

1 AAATTAATAC GACTCACTAT AGGCCTTCA GTGAGGGTCT GAGT

### 8.1.18 Frw.primers no (7)

---

1 AAATTAATAC GACTCACTAT AGGACCTGG ACAACCTCAA AGGCA

### 8.1.19 Frw.primers no (8)

---

1 AAATTAATAC GACTCACTAT AGGACCTTT GCTAAGCTGA GTGAA

## 8.1 Sequences of RNA Constructs

---

### 8.1.20 Frw.primers no (9)

---

1 AAATTAATAC GACTCACTAT AGGGACTGCA CTGTGACAAG CTGCA

### 8.1.21 Frw.primers no (10)

---

1 AAATTAATAC GACTCACTAT AGGGCGTGGA TCCTGAGAAC TTCA

### 8.1.22 Promoter 7

---

1 AAATTAATAC GACTCACTAT AGGG

### 8.1.23 $\beta$ -globin 5' Cy5

---

1 Cy5-UAGACA ACCAGCAGCC C-biotin

### 8.1.24 $\alpha$ -U1

---

1 GCCAGGUAAG UAU-biotin

## 8.1 Sequences of RNA Constructs

---

### 8.1.25 $\alpha$ -U6

---

1 CUGUGUAUCG UUCCAAUUUU

### 8.1.26 SRM160

---

1 MDAGFFRGTS AEQDNRFNSK QKKLLKQLKF AECLEKKVDM SKVNLEVIKP  
51 WITKRVTEIL GFEDDVVIEF IFNQLEVKNP DSKMMQINLT GFLNGKNARE  
101 FMGELWPLLL SAQENIAGIP SAFLELKKEE IKQRQIEQEK LASMKKQDED  
151 KDKRDKEEKE SSREKRERSR SPRRRKSRSP SPRRRSSPVR RERKRSHSR  
201 PRHRTKSRSP SPAPEKKEKT PELPEPSVKV KEPSVQEATS TSDILKVPKP  
251 EPIPEPKEPS PEKNSKKEKE KEKTRPRSR SRSRSRTRS RSPSHTRPRR  
301 RHRSRRSYS PRRRPSPRRR PSPRRRTPPR RMPPPPRHR SRSPVRRRRR  
351 SSASLSGSSS SSSSSRSRSP PKKPPKRTSS PPRKTRRLSP SASPPRRRHR  
401 PSPPATPPPK TRHSPTPQQS NRTRKSRVSV SPGRTSGKVT KHKGTEKRES  
451 PSPAPKPRKV ELSESEEDKG GKMAAADSVQ QRRQYRRQNQ QSSSDSGSSS  
501 SSEDERPQRS HVKNGEVGRR RRHSPRSAS PSPRKRQKET SPRGRRRRSP  
551 SPPPTRRRRS PSPAPPPRR RTPTPPRRR TPSPPRRRS PSPRRYSPI  
601 QRRYSPSPPP KRRTASPPPP PKRRASPSPP PKRRVSHSPP PKQRSSPVTK  
651 RRSPSLSSKH RKGSSPSRST REARSPQPNK RHSPSPRPRA PQTSSSPPPV  
701 RRGASSSPQR RQSPSPSTRP IRRVSRTPPEP KKIKKAASPS PQSVRRVSSS  
751 RSVGSPEPA AKKPPAPPSP VQSQSPSTNW SPAVPVKKAK SPTSPSPPPR  
801 NSDQEGGGKK KKKKKDKKHK KDKKHKHKHK HKKEKAVAAA AAAAVTPAAI  
851 AAATTTLAQE EPVAAPEPKK ETESEAEDNL DDLEKHLREK ALRSMRKAQV

## 8.1 Sequences of RNA Constructs

---

### 8.1.27 Double-Labelled RNA

---

```
1 GGGAGGAGGA CGGAGGACGG AGGACA1UCG CUGUCUGCGA GGGCCAGCUG  
51 UGGGGUGAG UACUCCCUCU C
```

## 8.1 Sequences of RNA Constructs

Table 8.1: Single Molecule Data shows the colocalisation percentages. #M is the the number of identified marker spots, #C is the number of colocalized spots and #CE is the colocalization percentage.

Experiments	#M	#C	#CE	1	2	3	4	5	>5	X
GloC-SMN2-CY5/ E (-ATP)	942	100	11	35	24	17	11	3	4	9
GloC-SMN2-CY5/ A (+ATP+AntiU6)	584	91	16	49	29	10	2	2	2	8
GloC-SMN2-CY5/ A (+ATP+AntiU6+AntiU1)	561	100	18	28	25	20	7	3	5	15
GloC-SMN2-CY5/ A (+ATP+AntiU6+AntiU2)	1126	100	9	59	27	7	5	0	0	2
GloC-SMN2-CY5/ E (-ATP+phostop)	302	50	17	42	20	6	4	2	2	26
GloC-SMN2-CY5/ E (-ATP+phostop+AntiU1)	597	50	8	64	18	12	2	0	0	4
GloC-SMN2-CY5/ E (-ATP+AntiU1)	1085	100	9	48	25	16	7	0	0	4
GloC-SMN2-CY5/ +ESE/ E (-ATP)	423	100	24	42	25	11	9	4	4	9
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6)	676	103	15	39	32	13	7	0	0	10
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6+AntiU1)	469	100	21	44	25	12	4	1	1	14
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6+AntiU2)	523	102	20	48	31	17	3	0	0	1
GloC-SMN2-CY5/ +ESE/ E (-ATP+phostop)	332	50	15	62	16	8	4	2	2	8
GloC-SMN2-CY5/ +ESE/ E (-ATP+phostop+AntiU1)	199	50	25	50	30	10	4	0	0	6
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU1)	615	100	16	55	24	12	6	1	1	2
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU2)	523	102	20	48	31	17	3	0	0	1
GloC-SMN2-CY5/ +ESE/ E (-ATP-Phostop+AntiU1)	615	100	16	55	24	12	6	1	1	2
GloC-SMN2-CY5/ -ESE/ A (+ATP+AntiU2)	1126	100	9	59	27	7	5	0	0	2
GloC-SMN2-CY5/ A (+ATP+AntiU1+AntiU2)	176	21	12	90	5	0	0	0	0	5
GloC-SMN2-CY5/ A (+AntiU1+AntiU2)	146	20	14	85	10	0	0	0	0	5
GloC-SMN2-CY5/ A (+AntiU6+AntiU1)	183	20	11	80	10	0	0	0	0	10
GloC-SMN2-CY5/ A (+AntiU6+AntiU2)	382	20	5	75	15	0	0	0	0	10

Continued on next page

## 8.1 Sequences of RNA Constructs

Experiments	#M	#C	#CE	1	2	3	4	5	>5	X
GloC-SMN2-CY5/ E (-ATP+AntiU1+AntiU2+AntiU6)	315	22	7	50	27	14	0	0	0	9
GloC-SMN2-CY5/ E (-ATP+AntiU1+AntiU2)	241	20	8	90	10	0	0	0	0	0
GloC-SMN2-CY5/ E (-ATP+AntiU1+AntiU6)	192	22	11	68	27	0	0	0	0	5
GloC-SMN2-CY5/ E (-ATP+AntiU2+AntiU6)	321	20	6	85	15	0	0	0	0	0
GloC-SMN2-CY5/ E (-ATP+AntiU2)	357	22	6	73	14	5	0	0	0	9
GloC-SMN2-CY5/ E (-ATP+AntiU6)	222	20	9	75	20	0	0	0	0	5
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6+AntiU1+AntiU2)	260	20	8	75	25	0	0	0	0	0
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU1+AntiU2)	257	24	9	63	33	0	0	0	0	4
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6+AntiU1)	244	20	8	60	30	5	0	0	0	5
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6+AntiU2)	105	21	20	52	33	10	0	0	0	5
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU1+AntiU2+AntiU6)	545	22	4	59	32	5	0	0	0	5
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU1+AntiU2)	193	20	10	75	25	0	0	0	0	0
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU1+AntiU6)	275	22	8	73	27	0	0	0	0	0
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU2+AntiU6)	387	20	5	75	20	0	0	0	0	5
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU2)	498	22	4	73	23	0	0	0	0	5
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU6)	392	21	5	62	38	0	0	0	0	0
GloC 16 E3-CY5/ E (-ATP+AntiU1).	836	100	12	48	23	12	5	0	0	12
GloC 16 E3-CY5/ E (-ATP)	836	100	12	48	23	12	5	0	0	12
GloC 16 E3-CY5/ A (+ATP+AntiU6)	486	100	21	57	29	7	1	0	0	6
GloC 16 E3-CY5/ A (+ATP+AntiU6+AntiU1)	785	100	13	61	24	9	2	0	0	4
GloC 16 E3-CY5/ A (+ATP+AntiU6+AntiU2)	768	100	13	63	25	6	0	0	0	6

Continued on next page

## 8.1 Sequences of RNA Constructs

Experiments	#M	#C	#CE	1	2	3	4	5	>5	X
GloC-SMN2-CY5/ -T+ESE/ E (-ATP+Phostop+AntiU1)	554	100	18	43	21	13	7	2	2	14
GloC-SMN2-CY5/ -T+ESE/ E (-ATP+Phostop-AntiU1)	612	100	16	51	22	12	6	1	1	8
GloC-SMN2-CY5/ -T+ESE/ E (-ATP-Phostop+AntiU1)	526	100	19	44	22	15	5	2	2	12
GloC-SMN2-CY5/ -T/ E (-ATP+Phostop+AntiU1)	1175	100	9	52	24	13	6	1	1	4
GloC-SMN2-CY5/ -T/ E (-ATP+Phostop-AntiU1)	1249	100	8	46	23	14	5	1	2	10
GloC-SMN2-CY5/ -T/ E (-ATP-Phostop+AntiU1)	771	100	13	52	14	9	5	0	0	20
GloC-SMN2-CY5/ -T/ E (-ATP)	895	100	11	37	25	17	8	3	5	8
GloC-SMN2-CY5/ -T/ A (+ATP+AntiU6)	971	100	10	49	38	8	0	0	0	5
GloC-SMN2-CY5/ -T/ A (+ATP+AntiU6+AntiU1)	590	100	17	53	26	12	2	0	0	7
GloC-SMN2-CY5/ -T/ A (+ATP+AntiU6+AntiU2)	722	100	14	63	22	7	2	0	0	6
GloC-SMN2-CY5/ -T/ E (-ATP+phostop)	1249	100	8	46	23	14	5	1	2	10
GloC-SMN2-CY5/ -T/ E (-ATP+phostop+AntiU1)	1175	100	9	52	24	13	6	1	1	4
GloC-SMN2-CY5/ -T/ E (-ATP+AntiU1)	771	100	13	52	14	9	5	0	0	20
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP)	692	100	14	46	28	12	6	4	4	4
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6)	405	100	25	47	25	12	4	0	1	11
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6+AntiU1)	579	100	17	50	30	12	4	0	0	4
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6+AntiU2)	469	100	21	44	34	13	1	0	0	8
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP+phostop)	612	100	16	51	22	12	6	1	1	8
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP+phostop+AntiU1)	554	100	18	43	21	13	7	2	2	14
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP+AntiU1)	526	100	19	44	22	15	5	2	2	12
GloC-SMN2-CY5/ -T/ A (+ATP+AntiU6)	428	100	23	40	30	16	3	1	1	10

Continued on next page

## 8.1 Sequences of RNA Constructs

Experiments	#M	#C	#CE	1	2	3	4	5	>5	X
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6)	633	100	16	34	31	11	5	1	1	18
Dimerization/ no RNA/ RiboA/ A (+ATP+Anti-U6)	110	21	19	60	57	5	0	0	0	0
Dimerization/ no RNA/ RiboA/ E (-ATP)	104	79	76	19	6	0	0	0	0	3
GloC (Spliced RNA)-CY5/ A (+ATP+Anti-U6)	428	100	23	65	22	6	1	0	0	6
GloC (Spliced RNA)-CY5/ SRm160-GFP / E (-ATP)	351	100	28	48	21	8	4	2	2	17
GloC-Cy5/ Abundance of SRm160 1:1/ A (+ATP+AntiU6)	2019	100	10	61	26	11	1	0	0	1
GloC-Cy5/ Abundance of SRm160 1:1/ E (-ATP)	1251	100	8	51	28	12	5	0	0	4
GloC-Cy5/ Abundance of SRm160 1:2/ A (+ATP+AntiU6)	873	100	8	65	26	8	1	0	0	0
GloC-Cy5/ Abundance of SRm160 1:2/ E (-ATP)	2154	100	5	57	27	11	5	0	0	4

## 8.1 Sequences of RNA Constructs

Table 8.2: Single Molecule Data shows the number molecules in each bleaching step. #M is the the number of identified marker spots, #C is the number of colocalized spots and #CE is the colocalization percentage.

Experiments	#M	#C	#CE	1	2	3	4	5	>5	X
GloC-SMN2-CY5/E (-ATP)	942	100	11	35	24	17	11	3	4	9
GloC-SMN2-CY5/ A (+ATP+AntiU6)	584	91	16	45	26	9	2	2	2	7
GloC-SMN2-CY5/ A (+ATP+AntiU6+AntiU1)	561	100	18	28	25	20	7	3	5	15
GloC-SMN2-CY5/ A (+ATP+AntiU6+AntiU2)	1126	100	9	59	27	7	5	0	0	2
GloC-SMN2-CY5/ E (-ATP+phostop)	302	50	17	21	10	3	2	1	1	13
GloC-SMN2-CY5/ E (-ATP+phostop+AntiU1)	597	50	8	32	9	6	1	0	0	2
GloC-SMN2-CY5/ E (-ATP+AntiU1)	1085	100	9	48	25	16	7	0	0	4
GloC-SMN2-CY5/ +ESE/ E (-ATP)	423	100	24	42	25	11	9	4	4	9
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6)	676	103	15	40	33	13	7	0	0	10
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6+AntiU1)	469	100	21	44	25	12	4	1	1	14
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6+AntiU2)	523	102	20	49	32	17	3	0	0	1
GloC-SMN2-CY5/ +ESE/ E (-ATP+phostop)	332	50	15	31	8	4	2	1	1	4
GloC-SMN2-CY5/ +ESE/ E (-ATP+phostop+AntiU1)	199	50	25	25	15	5	2	0	0	3
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU1)	615	100	16	55	24	12	6	1	1	2
GloC-SMN2-CY5/ A (+AntiU1+AntiU2)	146	20	14	17	2	0	0	0	0	1
GloC-SMN2-CY5/ A (+AntiU6+AntiU1)	183	20	11	16	2	0	0	0	0	2
GloC-SMN2-CY5/ A (+AntiU6+AntiU2)	382	20	5	15	3	0	0	0	0	2
GloC-SMN2-CY5/ E (-ATP+AntiU1+AntiU2+AntiU6)	315	22	7	11	6	3	0	0	0	2
GloC-SMN2-CY5/ E (-ATP+AntiU1+AntiU2)	241	20	8	18	2	0	0	0	0	0
GloC-SMN2-CY5/ E (-ATP+AntiU1+AntiU6)	192	22	11	15	6	0	0	0	0	1
GloC-SMN2-CY5/ E (-ATP+AntiU2+AntiU6)	321	20	6	17	3	0	0	0	0	0

Continued on next page

## 8.1 Sequences of RNA Constructs

Experiments	#M	#C	#CE	1	2	3	4	5	>5	X
GloC-SMN2-CY5/ E (-ATP+AntiU2)	357	22	6	16	3	1	0	0	0	2
GloC-SMN2-CY5/ E (-ATP+AntiU6)	222	20	9	15	4	0	0	0	0	1
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6+AntiU1+AntiU2)	260	20	8	15	5	0	0	0	0	0
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU1+AntiU2)	257	24	9	15	8	0	0	0	0	1
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6+AntiU1)	244	20	8	12	6	1	0	0	0	1
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU6+AntiU2)	105	21	20	11	7	2	0	0	0	1
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU1+AntiU2+AntiU6)	545	22	4	13	7	1	0	0	0	1
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU1+AntiU2)	193	20	10	15	5	0	0	0	0	0
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU1+AntiU6)	275	22	8	16	6	0	0	0	0	0
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU2+AntiU6)	387	20	5	15	4	0	0	0	0	1
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU2)	498	22	4	16	5	0	0	0	0	1
GloC-SMN2-CY5/ +ESE/ E (-ATP+AntiU6)	392	21	5	13	8	0	0	0	0	0
GloC-SMN2-CY5/ +ESE/ A (+ATP+AntiU2)	523	102	20	49	32	17	3	0	0	1
GloC-SMN2-CY5/ +ESE/ E (-ATP-Phostop+AntiU1)	615	100	16	55	24	12	6	1	1	2
GloC-SMN2-CY5/ -ESE/ A (+ATP+AntiU2)	1126	100	9	59	27	7	5	0	0	2
GloC-SMN2-CY5/ A (+ATP+AntiU1+AntiU2)	176	21	12	19	1	0	0	0	0	1
GloC 16 E3-CY5/ E (-ATP+AntiU1).	836	100	12	48	23	12	5	0	0	12
GloC 16 E3-CY5/ E (-ATP)	836	100	12	48	23	12	5	0	0	12
GloC 16 E3-CY5/ A (+ATP+AntiU6)	486	100	21	57	29	7	1	0	0	6
GloC 16 E3-CY5/ A (+ATP+AntiU6+AntiU1)	785	100	13	61	24	9	2	0	0	4
GloC 16 E3-CY5/ A (+ATP+AntiU6+AntiU2)	768	100	13	63	25	6	0	0	0	6

Continued on next page

## 8.1 Sequences of RNA Constructs

Experiments	#M	#C	#CE	1	2	3	4	5	>5	X
GloC-SMN2-CY5/ -T+ESE/ E (-ATP+Phostop+AntiU1)	554	100	18	43	21	13	7	2	2	14
GloC-SMN2-CY5/ -T+ESE/ E (-ATP+Phostop-AntiU1)	612	100	16	51	22	12	6	1	1	8
GloC-SMN2-CY5/ -T+ESE/ E (-ATP-Phostop+AntiU1)	526	100	19	44	22	15	5	2	2	12
GloC-SMN2-CY5/ -T/ E (-ATP+Phostop+AntiU1)	1175	100	9	52	24	13	6	1	1	4
GloC-SMN2-CY5/ -T/ E (-ATP+Phostop-AntiU1)	1249	100	8	46	23	14	5	1	2	10
GloC-SMN2-CY5/ -T/ E (-ATP-Phostop+AntiU1)	771	100	13	52	14	9	5	0	0	20
GloC-SMN2-CY5/ -T/ E (-ATP)	895	100	11	37	25	17	8	3	5	8
GloC-SMN2-CY5/ -T/ A (+ATP+AntiU6)	971	100	10	49	38	8	0	0	0	5
GloC-SMN2-CY5/ -T/ A (+ATP+AntiU6+AntiU1)	590	100	17	53	26	12	2	0	0	7
GloC-SMN2-CY5/ -T/ A (+ATP+AntiU6+AntiU2)	722	100	14	63	22	7	2	0	0	6
GloC-SMN2-CY5/ -T/ E (-ATP+phostop)	1249	100	8	46	23	14	5	1	2	10
GloC-SMN2-CY5/ -T/ E (-ATP+phostop+AntiU1)	1175	100	9	52	24	13	6	1	1	4
GloC-SMN2-CY5/ -T/ E (-ATP+AntiU1)	771	100	13	52	14	9	5	0	0	20
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP)	692	100	14	46	28	12	6	4	4	4
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6)	405	100	25	47	25	12	4	0	1	11
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6+AntiU1)	579	100	17	50	30	12	4	0	0	4
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6+AntiU2)	469	100	21	44	34	13	1	0	0	8
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP+phostop)	612	100	16	51	22	12	6	1	1	8
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP+phostop+AntiU1)	554	100	18	43	21	13	7	2	2	14
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP+AntiU1)	526	100	19	44	22	15	5	2	2	12
GloC-SMN2-CY5/-T/ A (+ATP+AntiU6)	428	100	23	40	30	16	3	1	1	10

Continued on next page

## 8.1 Sequences of RNA Constructs

Experiments	#M	#C	#CE	1	2	3	4	5	>5	X
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6)	633	100	16	34	31	11	5	1	1	18
Dimerization/ no RNA/ RiboA/ A (+ATP+Anti-U6)	110	21	19	13	12	1	0	0	0	0
Dimerization/ no RNA/ RiboA/ E (-ATP)	104	79	76	15	5	0	0	0	0	2
GloC (Spliced RNA)-CY5/ A (+ATP+Anti-U6)	428	100	23	65	22	6	1	0	0	6
GloC (Spliced RNA)-CY5/ SRm160-GFP / E (-ATP)	351	100	28	48	21	8	4	2	2	17
GloC-Cy5/ Abundance of SRm160 1:1/ A (+ATP+AntiU6)	2019	100	10	61	26	11	1	0	0	1
GloC-Cy5/ Abundance of SRm160 1:1/ E (-ATP)	1251	100	8	51	28	12	5	0	0	4
GloC-Cy5/ Abundance of SRm160 1:2/ A (+ATP+AntiU6)	873	100	8	65	26	8	1	0	0	0
GloC-Cy5/ Abundance of SRm160 1:2/ E (-ATP)	2154	100	5	57	27	11	5	0	0	4

## 8.1 Sequences of RNA Constructs

Table 8.3: Single Molecule Data shows the error bars values. #M is the the number of identified marker spots, #C is the number of colocalized spots and #CE is the colocalization percentage.

Experiments	1	2	3	4	5	>5	X
GloC-SMN2-CY5/E (-ATP)	0.616	0.513	0.434	0.350	0.058	0.090	0.303
GloC-SMN2-CY5/ A (+ATP + AntiU6)	1.152	0.893	0.534	0.095	0.095	0.095	0.472
GloC-SMN2-CY5/ A (+ATP + AntiU6 + AntiU1)	0.919	0.871	0.783	0.469	0.165	0.354	0.681
GloC-SMN2-CY5/ A (+ATP + AntiU6 + AntiU2)	0.664	0.456	0.146	0.088	0.000	0.000	0.022
GloC-SMN2-CY5/ E (-ATP + phostop)	1.991	1.431	0.803	0.658	0.309	0.309	1.614
GloC-SMN2-CY5/ E (-ATP + phostop + AntiU1)	1.266	0.700	0.574	0.079	0.000	0.000	0.224
GloC-SMN2-CY5/ E (-ATP + AntiU1)	0.624	0.455	0.366	0.157	0.000	0.000	0.068
GloC-SMN2-CY5/ +ESE/ E (-ATP)	1.454	1.147	0.774	0.702	0.445	0.445	0.702
GloC-SMN2-CY5/ +ESE/ A (+ATP + AntiU6)	0.895	0.817	0.521	0.384	0.000	0.000	0.458
GloC-SMN2-CY5/ +ESE/ A (+ATP + AntiU6 + AntiU1)	1.346	1.037	0.729	0.362	0.045	0.045	0.786
GloC-SMN2-CY5/ +ESE/ A (+ATP + AntiU6 + AntiU2)	1.263	1.038	0.768	0.184	0.000	0.000	0.035
GloC-SMN2-CY5/ +ESE/ E (-ATP + phostop)	2.139	1.175	0.842	0.599	0.256	0.256	0.842
GloC-SMN2-CY5/ +ESE/ E (-ATP + phostop + AntiU1)	3.075	2.536	1.549	0.995	0.000	0.000	1.212
GloC-SMN2-CY5/ +ESE/ E (-ATP + AntiU1)	1.151	0.781	0.558	0.387	0.026	0.026	0.075
GloC-SMN2-CY5/ A (+ATP + AntiU1 + AntiU2)	3.767	1.223	0.000	0.000	0.000	0.000	1.223
GloC-SMN2-CY5/ A (+AntiU1 + AntiU2)	4.082	2.090	0.000	0.000	0.000	0.000	1.505
GloC-SMN2-CY5/ A (+AntiU6 + AntiU1)	3.667	1.680	0.000	0.000	0.000	0.000	1.680
GloC-SMN2-CY5/ A (+AntiU6 + AntiU2)	2.032	0.994	0.000	0.000	0.000	0.000	0.817
GloC-SMN2-CY5/ E (-ATP + AntiU1 + AntiU2 + AntiU6)	2.059	1.584	1.147	0.000	0.000	0.000	0.943
GloC-SMN2-CY5/ E (-ATP + AntiU1 + AntiU2)	3.116	1.285	0.000	0.000	0.000	0.000	0.000
GloC-SMN2-CY5/ E (-ATP + AntiU1 + AntiU6)	3.454	2.519	0.000	0.000	0.000	0.000	1.097

Continued on next page

## 8.1 Sequences of RNA Constructs

Experiments	1	2	3	4	5	>5	X
GloC-SMN2-CY5/ E (-ATP + AntiU2 + AntiU6)	2.463	1.178	0.000	0.000	0.000	0.000	0.000
GloC-SMN2-CY5/ E (-ATP + AntiU2)	2.132	1.014	0.593	0.000	0.000	0.000	0.834
GloC-SMN2-CY5/ E (-ATP + AntiU6)	3.174	1.922	0.000	0.000	0.000	0.000	0.996
GloC-SMN2-CY5/ +ESE/ A (+ATP + AntiU6 + AntiU1 + AntiU2)	2.810	1.828	0.000	0.000	0.000	0.000	0.000
GloC-SMN2-CY5/ +ESE/ A (+ATP + AntiU1 + AntiU2)	2.676	2.096	0.000	0.000	0.000	0.000	0.788
GloC-SMN2-CY5/ +ESE/ A (+ATP + AntiU6 + AntiU1)	2.757	2.102	0.907	0.000	0.000	0.000	0.907
GloC-SMN2-CY5/ +ESE/ A (+ATP + AntiU6 + AntiU2)	4.879	4.543	2.803	0.000	0.000	0.000	2.031
GloC-SMN2-CY5/ +ESE/ E (-ATP + AntiU1 + AntiU2 + AntiU6)	1.332	1.004	0.325	0.000	0.000	0.000	0.325
GloC-SMN2-CY5/ +ESE/ E (-ATP + AntiU1 + AntiU2)	3.509	2.417	0.000	0.000	0.000	0.000	0.000
GloC-SMN2-CY5/ +ESE/ E (-ATP + AntiU1 + AntiU6)	2.660	1.802	0.000	0.000	0.000	0.000	0.000
GloC-SMN2-CY5/ +ESE/ E (-ATP + AntiU2 + AntiU6)	2.009	1.125	0.000	0.000	0.000	0.000	0.574
GloC-SMN2-CY5/ +ESE/ E (-ATP + AntiU2)	1.582	0.935	0.000	0.000	0.000	0.000	0.389
GloC-SMN2-CY5/ +ESE/ E (-ATP + AntiU6)	1.842	1.496	0.000	0.000	0.000	0.000	0.000
GloC-SMN2-CY5/ +ESE/ A (+ATP + AntiU2)	1.263	1.038	0.768	0.184	0.000	0.000	0.035
GloC-SMN2-CY5/ +ESE/ E (-ATP-Phostop + AntiU1)	1.151	0.781	0.558	0.387	0.026	0.026	0.075
GloC-SMN2-CY5/ -ESE/ A (+ATP + AntiU2)	0.664	0.456	0.146	0.088	0.000	0.000	0.022
GloC 16 E3-CY5/ E (-ATP + AntiU1).	0.805	0.566	0.411	0.159	0.000	0.000	0.411
GloC 16 E3-CY5/ E (-ATP)	0.805	0.566	0.411	0.159	0.000	0.000	0.411
GloC 16 E3-CY5/ A (+ATP + AntiU6)	1.460	1.074	0.540	0.042	0.000	0.000	0.501
GloC 16 E3-CY5/ A (+ATP + AntiU6 + AntiU1)	0.955	0.614	0.380	0.046	0.000	0.000	0.129
GloC 16 E3-CY5/ A (+ATP + AntiU6 + AntiU2)	0.990	0.640	0.248	0.000	0.000	0.000	0.248

Continued on next page

## 8.1 Sequences of RNA Constructs

Experiments	1	2	3	4	5	>5	X
GloC-SMN2-CY5/ -T+ESE/ E (-ATP+Phostop+AntiU1)	1.137	0.811	0.643	0.475	0.092	0.092	0.667
GloC-SMN2-CY5/ -T+ESE/ E (-ATP+Phostop-AntiU1)	1.117	0.753	0.560	0.390	0.027	0.027	0.459
GloC-SMN2-CY5/ -T+ESE/ E (-ATP-Phostop+AntiU1)	1.207	0.873	0.726	0.402	0.102	0.102	0.651
GloC-SMN2-CY5/ -T/ E (-ATP+Phostop+AntiU1)	0.600	0.413	0.305	0.106	0.007	0.007	0.058
GloC-SMN2-CY5/ -T/ E (-ATP+Phostop-AntiU1)	0.533	0.380	0.298	0.072	0.006	0.018	0.202
GloC-SMN2-CY5/ -T/ E (-ATP-Phostop+AntiU1)	0.903	0.481	0.387	0.187	0.000	0.000	0.572
GloC-SMN2-CY5/ -T/ E (-ATP)	0.665	0.551	0.456	0.281	0.065	0.139	0.281
GloC-SMN2-CY5/ -T/ A (+ATP+AntiU6)	0.702	0.622	0.239	0.000	0.000	0.000	0.118
GloC-SMN2-CY5/ -T/ A (+ATP+AntiU6+AntiU1)	1.177	0.845	0.581	0.081	0.000	0.000	0.446
GloC-SMN2-CY5/ -T/ A (+ATP+AntiU6+AntiU2)	1.050	0.640	0.354	0.054	0.000	0.000	0.281
GloC-SMN2-CY5/ -T/ E (-ATP+phostop)	0.533	0.380	0.298	0.072	0.006	0.018	0.202
GloC-SMN2-CY5/ -T/ E (-ATP+phostop+AntiU1)	0.600	0.413	0.305	0.106	0.007	0.007	0.058
GloC-SMN2-CY5/ -T/ E (-ATP+AntiU1)	0.903	0.481	0.387	0.187	0.000	0.000	0.572
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP)	0.947	0.749	0.496	0.306	0.167	0.167	0.167
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6)	1.592	1.196	0.843	0.485	0.000	0.061	0.808
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6+AntiU1)	1.167	0.921	0.592	0.238	0.000	0.000	0.238
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6+AntiU2)	1.346	1.197	0.758	0.045	0.000	0.000	0.598
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP+phostop)	1.117	0.753	0.560	0.390	0.027	0.027	0.459
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP+phostop+AntiU1)	1.137	0.811	0.643	0.475	0.092	0.092	0.667
GloC-SMN2-CY5/ -T/ +ESE/ E (-ATP+AntiU1)	1.207	0.873	0.726	0.402	0.102	0.102	0.651
GloC-SMN2-CY5/-T/ A (+ATP+AntiU6)	1.407	1.234	0.917	0.283	0.055	0.055	0.730

Continued on next page

## 8.1 Sequences of RNA Constructs

Experiments	1	2	3	4	5	>5	X
GloC-SMN2-CY5/ -T/ +ESE/ A (+ATP+AntiU6)	0.896	0.858	0.519	0.278	0.025	0.025	0.661
Dimerization/ no RNA/ RiboA/ A (+ATP+Anti-U6)	4.748	4.764	1.940	0.000	0.000	0.000	0.000
Dimerization/ no RNA/ RiboA/ E (-ATP)	3.807	2.344	0.000	0.000	0.000	0.000	1.511
GloC (Spliced RNA)-CY5/ A (+ATP+Anti-U6)	1.735	1.067	0.568	0.055	0.000	0.000	0.568
GloC (Spliced RNA)-CY5/ SRm160-GFP / E (-ATP)	1.834	1.266	0.797	0.567	0.229	0.229	1.146
GloC-Cy5/ Abundance of SRm160 1:1/ A (+ATP+AntiU6)	0.381	0.251	0.089	0.002	0.000	0.000	0.002
GloC-Cy5/ Abundance of SRm160 1:1/ E (-ATP)	0.559	0.418	0.264	0.071	0.000	0.000	0.051
GloC-Cy5/ Abundance of SRm160 1:2/ A (+ATP+AntiU6)	0.888	0.575	0.296	0.013	0.000	0.000	0.000
GloC-Cy5/ Abundance of SRm160 1:2/ E (-ATP)	0.346	0.240	0.078	0.024	0.000	0.000	0.017

# Bibliography

- Abelson, J., Blanco, M., Ditzler, M. A., Fuller, F., Aravamudhan, P., Wood, M., Villa, T., Ryan, D. E., Pleiss, J. A., Maeder, C., et al. (2010). Conformational dynamics of single pre-mrna molecules during in vitro splicing. *Nature structural & molecular biology*, 17(4):504–512.
- Aebi, M., Hornig, H., Padgett, R. A., Reiser, J., and Weissmann, C. (1986). Sequence requirements for splicing of higher eukaryotic nuclear pre-mrna. *Cell*, 47(4):555–565.
- Aebi, M., Hornig, H., and Weissmann, C. (1987). 5' cleavage site in eukaryotic pre-mrna splicing is determined by the overall 5' splice region, not by the conserved 5' GU. *Cell*, 50(2):237–246.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002). Cell junctions, cell adhesion, and the extracellular matrix.
- Amrein, H., Gorman, M., and Nöthiger, R. (1988). The sex-determining gene tra-2 of drosophila encodes a putative rna binding protein. *Cell*, 55(6):1025–1035.
- Amrein, H., Hedley, M. L., and Maniatis, T. (1994). The role of specific

## BIBLIOGRAPHY

---

- protein-rna and protein-protein interactions in positive and negative control of pre-mrna splicing by transformer 2. *Cell*, 76(4):735–746.
- Ars, E., Serra, E., García, J., Krüyer, H., Gaona, A., Lázaro, C., and Estivill, X. (2000). Mutations affecting mrna splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Human molecular genetics*, 9(2):237–247.
- Axelrod, D. (2001). Total internal reflection fluorescence microscopy in cell biology. *Traffic*, 2(11):764–774.
- Baltimore, D. (1970). Viral rna-dependent dna polymerase: Rna-dependent dna polymerase in virions of rna tumour viruses. *Nature*, 226(5252):1209–1211.
- Barabino, S. M., Blencowe, B. J., Ryder, U., Sproat, B. S., and Lamond, A. I. (1990). Targeted snrnp depletion reveals an additional role for mammalian u1 snrnp in spliceosome assembly. *Cell*, 63(2):293–302.
- Barbosa-Morais, N. L., Irimia, M., Pan, Q., Xiong, H. Y., Gueroussov, S., Lee, L. J., Slobodeniuc, V., Kutter, C., Watt, S., Çolak, R., et al. (2012). The evolutionary landscape of alternative splicing in vertebrate species. *Science*, 338(6114):1587–1593.
- Bell, M. V., Cowper, A. E., Lefranc, M.-P., Bell, J. I., and Sreaton, G. R. (1998). Influence of intron length on alternative splicing of cd44. *Molecular and Cellular Biology*, 18(10):5930–5941.
- Bennett, M., Michaud, S., Kingston, J., and Reed, R. (1992). Protein compo-

## BIBLIOGRAPHY

---

- nents specifically associated with prespliceosome and spliceosome complexes. *Genes & Development*, 6(10):1986–2000.
- Bentley, D. (2002). The mrna assembly line: transcription and processing machines in the same factory. *Current opinion in cell biology*, 14(3):336–342.
- Bentley, D. L. (2014). Coupling mrna processing with transcription in time and space. *Nature Reviews Genetics*, 15(3):163–175.
- Berezin, M. Y. and Achilefu, S. (2010). Fluorescence lifetime measurements and biological imaging. *Chemical reviews*, 110(5):2641–2684.
- Berget, S. M. (1995). Exon recognition in vertebrate splicing. *Journal of biological Chemistry*, 270(6):2411–2414.
- Berget, S. M., Moore, C., and Sharp, P. A. (1977). Spliced segments at the 5 terminus of adenovirus 2 late mrna. *Proceedings of the National Academy of Sciences*, 74(8):3171–3175.
- 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):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):781–787.
- Berlier, J. E., Rothe, A., Buller, G., Bradford, J., Gray, D. R., Filanoski, B. J., Telford, W. G., Yue, S., Liu, J., Cheung, C.-Y., et al. (2003). Quantitative comparison of long-wavelength alexa fluor dyes to cy dyes: fluorescence of

## BIBLIOGRAPHY

---

- the dyes and their bioconjugates. *Journal of Histochemistry & Cytochemistry*, 51(12):1699–1712.
- Bessonov, S., Anokhina, M., Krasauskas, A., Golas, M. M., Sander, B., Will, C. L., Urlaub, H., Stark, H., and Lührmann, R. (2010). Characterization of purified human bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. *Rna*.
- Bindereif, A. and Green, M. R. (1987). An ordered pathway of snrnp binding during mammalian pre-mrna splicing complex assembly. *The EMBO journal*, 6(8):2415.
- Bingaman, S., Huxley, V. H., and Rumbaut, R. E. (2003). Fluorescent dyes modify properties of proteins used in microvascular research. *Microcirculation*, 10(2):221–231.
- Birney, E., Kumar, S., and Krainer, A. R. (1993). Analysis of the rna-recognition motif and rs and rgg domains: conservation in metazoan pre-mrna splicing factors. *Nucleic acids research*, 21(25):5803–5816.
- Black, D. L. (2003). Mechanisms of alternative pre-messenger rna splicing. *Annual review of biochemistry*, 72(1):291–336.
- 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):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):697–704.

## BIBLIOGRAPHY

---

- 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):1939–1952.
- Blencowe, B. J. (2000). Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends in biochemical sciences*, 25(3):106–110.
- Blencowe, B. J., Baurén, G., Eldridge, A. G., Issner, R., Nickerson, J. A., Rosonina, E., and Sharp, P. A. (2000). The srm160/300 splicing coactivator subunits. *Rna*, 6(01):111–120.
- Blencowe, B. J., Bowman, J. A., McCracken, S., and Rosonina, E. (1999). Sr-related proteins and the processing of messenger rna precursors. *Biochemistry and cell biology*, 77(4):277–291.
- Blencowe, B. J., Issner, R., Nickerson, J. A., and Sharp, P. A. (1998). A coactivator of pre-mrna splicing. *Genes & development*, 12(7):996–1009.
- Blencowe, B. J., Nickerson, J. A., Issner, R., Penman, S., and Sharp, P. A. (1994). Association of nuclear matrix antigens with exon-containing splicing complexes. *The Journal of Cell Biology*, 127(3):593–607.
- Brack, C. and Tonegawa, S. (1977). Variable and constant parts of the immunoglobulin light chain gene of a mouse myeloma cell are 1250 non-translated bases apart. *Proceedings of the National Academy of Sciences*, 74(12):5652–5656.
- Breathnach, R., Benoist, C., O'hare, K., Gannon, F., and Chambon, P. (1978). Ovalbumin gene: evidence for a leader sequence in mrna and dna sequences

## BIBLIOGRAPHY

---

- at the exon-intron boundaries. *Proceedings of the National Academy of Sciences*, 75(10):4853–4857.
- Brown, S. J., Stoilov, P., and Xing, Y. (2012). Chromatin and epigenetic regulation of pre-mrna processing. *Human molecular genetics*, page dds353.
- Buée, L., Bussiere, T., Buée-Scherrer, V., Delacourte, A., and Hof, P. R. (2000). Tau protein isoforms, phosphorylation and role in neurodegenerative disorders. *Brain Research Reviews*, 33(1):95–130.
- Buratti, E. and Baralle, F. E. (2004). Influence of rna secondary structure on the pre-mrna splicing process. *Molecular and cellular biology*, 24(24):10505–10514.
- Buratti, E., Dhir, A., Lewandowska, M. A., and Baralle, F. E. (2007). Rna structure is a key regulatory element in pathological atm and cftr pseudoexon inclusion events. *Nucleic acids research*, 35(13):4369–4383.
- Burge, C. B., Tuschl, T., and Sharp, P. A. (1999). 20 splicing of precursors to mRNAs by the spliceosomes. *Cold Spring Harbor Monograph Archive*, 37:525–560.
- Buschmann, V., Weston, K. D., and Sauer, M. (2003). Spectroscopic study and evaluation of red-absorbing fluorescent dyes. *Bioconjugate chemistry*, 14(1):195–204.
- Buvoli, M., Cobianchi, F., Biamonti, G., and Riva, S. (1990). Recombinant hnRNP protein A1 and its N-terminal domain show preferential affinity for oligodeoxynucleotides homologous to intron/exon acceptor sites. *Nucleic acids research*, 18(22):6595–6600.

## BIBLIOGRAPHY

---

- Buvoli, M., Cobianchi, F., and Riva, S. (1992). Interaction of hnrnp a1 with snrnps and pre-mrnas: evidence for a possible role of a1 rna annealing activity in the first steps of spliceosome assembly. *Nucleic acids research*, 20(19):5017–5025.
- Cáceres, J. F., Stamm, S., Helfman, D. M., Krainer, A. R., et al. (1994). Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *SCIENCE-NEW YORK THEN WASHINGTON-*, pages 1706–1706.
- Cao, W., Jamison, S. F., and Garcia-Blanco, M. A. (1997). Both phosphorylation and dephosphorylation of asf/sf2 are required for pre-mrna splicing in vitro. *Rna*, 3(12):1456–1467.
- Carmel, I., Tal, S., Vig, I., and Ast, G. (2004). Comparative analysis detects dependencies among the 5 splice-site positions. *Rna*, 10(5):828–840.
- Cartegni, L., Chew, S. L., and Krainer, A. R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nature Reviews Genetics*, 3(4):285–298.
- Cartegni, L. and Krainer, A. R. (2002). Disruption of an sf2/asf-dependent exonic splicing enhancer in smn2 causes spinal muscular atrophy in the absence of smn1. *Nature genetics*, 30(4):377–384.
- Casanova, D., Giaume, D., Moreau, M., Martin, J.-L., Gacoin, T., Boilot, J.-P., and Alexandrou, A. (2007). Counting the number of proteins coupled to single nanoparticles. *Journal of the American Chemical Society*, 129(42):12592–12593.

## BIBLIOGRAPHY

---

- Chabot, B., Black, D. L., LeMaster, D. M., and Steitz, J. A. (1985). The 3'splice site of pre-messenger rna is recognized by small nuclear.
- Chalfie, M. (1994). Green fluorescent protein as a marker for gene expression. *Trends in Genetics*, 10(5):151.
- Cheng, S. and Abelson, J. (1987). Spliceosome assembly in yeast. *Genes & Development*, 1(9):1014–1027.
- 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):2161–2172.
- Chiara, M. D., Gozani, O., Bennett, M., Champion-Arnaud, P., Palandjian, L., and Reed, R. (1996). Identification of proteins that interact with exon sequences, splice sites, and the branchpoint sequence during each stage of spliceosome assembly. *Molecular and Cellular Biology*, 16(7):3317–3326.
- 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*, 108(20):8233–8238.
- Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J., Bell, J., and Duncan, P. (1996). The clk/sty protein kinase phosphorylates sr splicing factors and regulates their intranuclear distribution. *The EMBO journal*, 15(2):265.

## BIBLIOGRAPHY

---

- 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*, 14(1):170–179.
- Crawford, D. J., Hoskins, A. A., Friedman, L. J., Gelles, J., and Moore, M. J. (2013). Single-molecule colocalization fret evidence that spliceosome activation precedes stable approach of 5 splice site and branch site. *Proceedings of the National Academy of Sciences*, 110(17):6783–6788.
- Créton, R. and Jaffe, L. F. (2001). Chemiluminescence microscopy as a tool in biomedical research. *Biotechniques*, 31(5):1098–1105.
- Crick, F. et al. (1970). Central dogma of molecular biology. *Nature*, 227(5258):561–563.
- 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 THEN WASHINGTON-*, pages 1866–1866.
- Darnell, J. E. (2013). Reflections on the history of pre-mrna processing and highlights of current knowledge: a unified picture. *Rna*, 19(4):443–460.
- David, C. J. and Manley, J. L. (2008). The search for alternative splicing regulators: new approaches offer a path to a splicing code. *Genes & development*, 22(3):279–285.
- Day, R. N. and Davidson, M. W. (2009). The fluorescent protein palette: tools for cellular imaging. *Chemical Society Reviews*, 38(10):2887–2921.
- De Conti, L., Baralle, M., and Buratti, E. (2013). Exon and intron definition in pre-mrna splicing. *Wiley Interdisciplinary Reviews: RNA*, 4(1):49–60.

## BIBLIOGRAPHY

---

- Deckert, J., Hartmuth, K., Boehringer, D., Behzadnia, N., Will, C. L., Kastner, B., Stark, H., Urlaub, H., and Lührmann, 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):5528–5543.
- Dempsey, G. T., Wang, W., and Zhuang, X. (2009). Fluorescence imaging at sub-diffraction-limit resolution with stochastic optical reconstruction microscopy. In *Handbook of single-molecule biophysics*, pages 95–127. Springer.
- Dönmez, G., Hartmuth, K., and Lührmann, R. (2004). Modified nucleotides at the 5 end of human u2 snrna are required for spliceosomal e-complex formation. *Rna*, 10(12):1925–1933.
- Dreyfuss, G., Kim, V. N., and Kataoka, N. (2002). Messenger-rna-binding proteins and the messages they carry. *Nature reviews Molecular cell biology*, 3(3):195–205.
- Duncan, P. I., Stojdl, D. F., Marius, R. M., and Bell, J. C. (1997). In vivo regulation of alternative pre-mrna splicing by the clk1 protein kinase. *Molecular and cellular biology*, 17(10):5996–6001.
- 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*, 96(11):6125–6130.
- Eperon, I., Ireland, D., Smith, R., Mayeda, A., and Krainer, A. (1993). Pathways for selection of 5′splice sites by u1 snrnps and sf2/asf. *The EMBO journal*, 12(9):3607.

## BIBLIOGRAPHY

---

- Eperon, I. C., Makarova, O. V., Mayeda, A., Munroe, S. H., Cáceres, 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):8303–8318.
- Eperon, L., Estibeiro, J., and Eperon, I. (1986). The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger rna.
- Erkelenz, S., Mueller, W. F., Evans, M. S., Busch, A., Schöneweis, K., Hertel, K. J., and Schaal, H. (2013). Position-dependent splicing activation and repression by sr and hnrnp proteins rely on common mechanisms. *Rna*, 19(1):96–102.
- Expert-Bezançon, A., Sureau, A., Durosay, P., Salesse, R., Groeneveld, H., Lecaer, J. P., and Marie, J. (2004). hnrnp a1 and the sr proteins asf/sf2 and sc35 have antagonistic functions in splicing of  $\beta$ -tropomyosin exon 6b. *Journal of Biological Chemistry*, 279(37):38249–38259.
- Fairbrother, W. G. and Chasin, L. A. (2000). Human genomic sequences that inhibit splicing. *Molecular and Cellular Biology*, 20(18):6816–6825.
- Fei, Y., Sun, Y.-S., Li, Y., Lau, K., Yu, H., Chokhawala, H. A., Huang, S., Landry, J. P., Chen, X., and Zhu, X. (2011). Fluorescent labeling agents change binding profiles of glycan-binding proteins. *Molecular BioSystems*, 7(12):3343–3352.
- Fetzer, S., Lauber, J., Will, C. L., and Lührmann, R. (1997). The [u4/u6. u5] tri-snrnp-specific 27k protein is a novel sr protein that can be phosphorylated by the snrnp-associated protein kinase. *Rna*, 3(4):344–355.

## BIBLIOGRAPHY

---

- Fialcowitz, E. J., Brewer, B. Y., Keenan, B. P., and Wilson, G. M. (2005). A hairpin-like structure within an au-rich mrna-destabilizing element regulates trans-factor binding selectivity and mrna decay kinetics. *Journal of Biological Chemistry*, 280(23):22406–22417.
- Fisette, J.-F., Toutant, J., Dugré-Brisson, S., Desgroseillers, L., and Chabot, B. (2010). hnrnp a1 and hnrnp h can collaborate to modulate 5 splice site selection. *Rna*, 16(1):228–238.
- Flaherty, S. M., Fortes, P., Izaurralde, E., Mattaj, I. W., and Gilmartin, G. M. (1997). Participation of the nuclear cap binding complex in pre-mrna 3 processing. *Proceedings of the National Academy of Sciences*, 94(22):11893–11898.
- Forkey, J. N., Quinlan, M. E., Shaw, M. A., Corrie, J. E., and Goldman, Y. E. (2003). Three-dimensional structural dynamics of myosin v by single-molecule fluorescence polarization. *Nature*, 422(6930):399–404.
- Förster, T. (1955). *Intermolecular energy transfer and fluorescence*. National Research Council of Canada.
- Fortes, P., Inada, T., Preiss, T., Hentze, M. W., Mattaj, I. W., and Sachs, A. B. (2000). The yeast nuclear cap binding complex can interact with translation factor eif4g and mediate translation initiation. *Molecular cell*, 6(1):191–196.
- Friedman, L. J. and Gelles, J. (2012). Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. *Cell*, 148(4):679–689.

## BIBLIOGRAPHY

---

- Fu, X.-D. (1995). The superfamily of arginine/serine-rich splicing factors. *Rna*, 1(7):663.
- 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*, 89(23):11224–11228.
- Fuda, N. J., Ardehali, M. B., and Lis, J. T. (2009). Defining mechanisms that regulate rna polymerase ii transcription in vivo. *Nature*, 461(7261):186–192.
- Gaur, R. K., Valcarcel, J., and Green, M. R. (1995). Sequential recognition of the pre-mrna branch point by u2af65 and a novel spliceosome-associated 28-kda protein. *Rna*, 1(4):407–417.
- 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):25–34.
- Gilbert, W. (1978). Why genes in pieces? *Nature*, 271(5645):501.
- Gooding, C., Roberts, G. C., and Smith, C. W. (1998). Role of an inhibitory pyrimidine element and polypyrimidine tract binding protein in repression of a regulated [alpha]-tropomyosin exon. *Rna*, 4(01):85–100.
- Goralski, T. J., Edström, J.-E., and Baker, B. S. (1989). The sex determination locus transformer-2 of drosophila encodes a polypeptide with similarity to rna binding proteins. *Cell*, 56(6):1011–1018.
- Görnemann, 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):53–63.

## BIBLIOGRAPHY

---

- Goarani, O., Potashkin, J., and Reed, R. (1998). A potential role for u2af-sap155 interactions in recruiting u2 snrnp to the branch site. *Molecular and cellular biology*, 18(8):4752–4760.
- Grabowski, P. J. and Sharp, P. A. (1986). Affinity chromatography of splicing complexes: U2, u5, and u4+ u6 small nuclear ribonucleoprotein particles in the spliceosome. *Science*, 233(4770):1294–1299.
- Graewe, S., Retzlaff, S., Struck, N., Janse, C. J., and Heussler, V. T. (2009). Going live: a comparative analysis of the suitability of the rfp derivatives redstar, mcherry and tdtomato for intravital and in vitro live imaging of plasmodium parasites. *Biotechnology journal*, 4(6):895–902.
- Graveley, B. R. (2000). Sorting out the complexity of sr protein functions. *Rna*, 6(09):1197–1211.
- Graveley, B. R. (2001). Alternative splicing: increasing diversity in the proteomic world. *TRENDS in Genetics*, 17(2):100–107.
- Graveley, B. R., Hertel, K. J., and Maniatis, T. (1998). A systematic analysis of the factors that determine the strength of pre-mrna splicing enhancers. *The EMBO journal*, 17(22):6747–6756.
- Graveley, B. R., Hertel, K. J., and Maniatis, T. (2001). The role of u2af35 and u2af65 in enhancer-dependent splicing. *Rna*, 7(6):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):765–771.

## BIBLIOGRAPHY

---

- Green, M. R. (1991). Biochemical mechanisms of constitutive and regulated pre-mrna splicing. *Annual review of cell biology*, 7(1):559–599.
- Grover, A., Houlden, H., Baker, M., Adamson, J., Lewis, J., Prihar, G., Pickering-Brown, S., Duff, K., and Hutton, M. (1999). 5 splice site mutations in tau associated with the inherited dementia ftdp-17 affect a stem-loop structure that regulates alternative splicing of exon 10. *Journal of Biological Chemistry*, 274(21):15134–15143.
- Gui, J.-F., Lane, W. S., and Fu, X.-D. (1994). A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature*, 369(6482):678–682.
- Günzl, A., Palfi, Z., and Bindereif, A. (2002). Analysis of rna–protein complexes by oligonucleotide-targeted rnase h digestion. *Methods*, 26(2):162–169.
- Han, J., Ding, J.-H., Byeon, C. W., Kim, J. H., Hertel, K. J., Jeong, S., and Fu, X.-D. (2011). Sr proteins induce alternative exon skipping through their activities on the flanking constitutive exons. *Molecular and cellular biology*, 31(4):793–802.
- Han, S. P., Tang, Y. H., and Smith, R. (2010). Functional diversity of the hnRNPs: past, present and perspectives. *Biochemical Journal*, 430(3):379–392.
- Hassler, K., Anhut, T., Rigler, R., Gösch, M., and Lasser, T. (2005). High count rates with total internal reflection fluorescence correlation spectroscopy. *Biophysical journal*, 88(1):L01–L03.

## BIBLIOGRAPHY

---

- Hastings, M. L. and Krainer, A. R. (2001). Pre-mrna splicing in the new millennium. *Current opinion in cell biology*, 13(3):302–309.
- Haugland, R. P. (2002). *Handbook of fluorescent probes and research products*. Molecular Probes.
- 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):1878–1886.
- Hodson, M. (2011). Using of sigle molecule methods to reveal the mechanisms of splice site selection. *Doctor of Philosophy*.
- Hodson, M. J., Hudson, A. J., Cherny, D., and Eperon, I. C. (2012). The transition in spliceosome assembly from complex e to complex a purges surplus u1 snrnps from alternative splice sites. *Nucleic acids research*, page gks322.
- Horowitz, D. S. and Krainer, A. R. (1994). Mechanisms for selecting 5 splice sites in mammalian pre-mrna splicing. *Trends in Genetics*, 10(3):100–106.
- 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. (2011a). Ordered and dynamic assembly of single spliceosomes. *Science*, 331(6022):1289–1295.
- Hoskins, A. A., Gelles, J., and Moore, M. J. (2011b). New insights into the spliceosome by single molecule fluorescence microscopy. *Current opinion in chemical biology*, 15(6):864–870.

## BIBLIOGRAPHY

---

- Hoskins, A. A. and Moore, M. J. (2012). The spliceosome: a flexible, reversible macromolecular machine. *Trends in biochemical sciences*, 37(5):179–188.
- Huranová, M., Ivani, I., Benda, A., Poser, I., Brody, Y., Hof, M., Shav-Tal, Y., Neugebauer, K. M., and Staněk, D. (2010). The differential interaction of snrnps with pre-mrna reveals splicing kinetics in living cells. *The Journal of cell biology*, 191(1):75–86.
- Hutton, M., Lendon, C. L., Rizzu, P., Baker, M., Froelich, S., Houlden, H., Pickering-Brown, S., Chakraverty, S., Isaacs, A., Grover, A., et al. (1998). Association of missense and 5-splice-site mutations in tau with the inherited dementia ftdp-17. *Nature*, 393(6686):702–705.
- Izaurrealde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I. W. (1994). A nuclear cap binding protein complex involved in pre-mrna splicing. *Cell*, 78(4):657–668.
- 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*, 89(12):5482–5486.
- Jeffreys, A. J. and Flavell, R. A. (1977). The rabbit  $\beta$ -globin gene contains a large insert in the coding sequence. *Cell*, 12(4):1097–1108.
- Juan, W. C., Roca, X., and Ong, S. T. (2014). Identification of cis-acting elements and splicing factors involved in the regulation of bim pre-mrna splicing. *PloS one*, 9(4):e95210.
- Jurica, M. S., Licklider, L. J., GYGI, S. P., Grigorieff, N., and Moore, M. J.

## BIBLIOGRAPHY

---

- (2002). Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *Rna*, 8(04):426–439.
- Kafasla, P., Mickleburgh, I., Llorian, M., Coelho, M., Gooding, C., Cherny, D., Joshi, A., Kotik-Kogan, O., Curry, S., Eperon, I. C., et al. (2012). Defining the roles and interactions of ptb. *Biochemical Society Transactions*, 40(4):815–820.
- 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):462–471.
- Kapanidis, A. N., Laurence, T. A., Lee, N. K., Margeat, E., Kong, X., and Weiss, S. (2005). Alternating-laser excitation of single molecules. *Accounts of chemical research*, 38(7):523–533.
- Karlström, A. and Nygren, P.-Å. (2001). Dual labeling of a binding protein allows for specific fluorescence detection of native protein. *Analytical biochemistry*, 295(1):22–30.
- Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M., and Stamm, S. (2013). Function of alternative splicing. *Gene*, 514(1):1–30.
- Kielkopf, C. L., Lücke, S., and Green, M. R. (2004). U2af homology motifs: protein recognition in the rrm world. *Genes & development*, 18(13):1513–1526.
- Kohtz, J. D., Jamison, S. F., Will, C. L., Zuo, P., Lührmann, R., Garcia-Blanco, M. A., and Manley, J. L. (1994). Protein–protein interactions and 5′-splice-site recognition in mammalian mrna precursors.

## BIBLIOGRAPHY

---

- 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):552–557.
- Konarska, M. M. and Sharp, P. A. (1986). Electrophoretic separation of complexes involved in the splicing of precursors to mrnas. *Cell*, 46(6):845–855.
- Konarska, M. M. and Sharp, P. A. (1987). Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell*, 49(6):763–774.
- Kornblihtt, A. R., de la Mata, M., Fededa, J. P., Munoz, M. J., and Nogues, G. (2004). Multiple links between transcription and splicing. *Rna*, 10(10):1489–1498.
- 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):35–42.
- Krainer, A. R. and Maniatis, T. (1985). Multiple factors including the small nuclear ribonucleoproteins u1 and u2 are necessary for pre-mrna splicing in vitro. *Cell*, 42(3):725–736.
- Krämer, A., Keller, W., Appel, B., and Lührmann, 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):299–307.
- Krämer, 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):1503.

## BIBLIOGRAPHY

---

- Kubin, R. F. and Fletcher, A. N. (1983). Fluorescence quantum yields of some rhodamine dyes. *Journal of Luminescence*, 27(4):455–462.
- Kuehner, J. N., Pearson, E. L., and Moore, C. (2011). Unravelling the means to an end: Rna polymerase ii transcription termination. *Nature reviews Molecular cell biology*, 12(5):283–294.
- Kwak, H. and Lis, J. T. (2013). Control of transcriptional elongation. *Annual review of genetics*, 47:483.
- Lacadie, S. A. and Rosbash, M. (2005). Cotranscriptional spliceosome assembly dynamics and the role of u1 snrna: 5 ss base pairing in yeast. *Molecular cell*, 19(1):65–75.
- Ladomery, M. (2013). Aberrant alternative splicing is another hallmark of cancer. *International journal of cell biology*, 2013.
- Lakowicz, J. (1999). Principles of fluorescence microscopy.
- Lakowicz, J. R. (2013). *Principles of fluorescence spectroscopy*. Springer Science & Business Media.
- Lamichhane, R., Daubner, G. M., Thomas-Crusells, J., Auweter, S. D., Manatschal, C., Austin, K. S., Valniuk, O., Allain, F. H.-T., and Rueda, D. (2010). Rna looping by ptb: evidence using fret and nmr spectroscopy for a role in splicing repression. *Proceedings of the National Academy of Sciences*, 107(9):4105–4110.
- Lamond, A. I., Konarska, M. M., Grabowski, P. J., and Sharp, P. A. (1988). Spliceosome assembly involves the binding and release of u4 small nu-

## BIBLIOGRAPHY

---

- clear ribonucleoprotein. *Proceedings of the National Academy of Sciences*, 85(2):411–415.
- Lamond, A. I., Konarska, M. M., and Sharp, P. A. (1987). A mutational analysis of spliceosome assembly: evidence for splice site collaboration during spliceosome formation. *Genes & Development*, 1(6):532–543.
- Landgraf, D., Okumus, B., Chien, P., Baker, T. A., and Paulsson, J. (2012). Segregation of molecules at cell division reveals native protein localization. *Nature methods*, 9(5):480–482.
- Langford, C. J., Klinz, F.-J., Donath, C., and Gallwitz, D. (1984). Point mutations identify the conserved, intron-contained tactaac box as an essential splicing signal sequence in yeast. *Cell*, 36(3):645–653.
- Lavigueur, A., La Branche, H., Kornblihtt, A. R., and Chabot, B. (1993). A splicing enhancer in the human fibronectin alternate ed1 exon interacts with sr proteins and stimulates u2 snrnp binding. *Genes & Development*, 7(12a):2405–2417.
- 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):355–358.
- Lear, A., Eperon, L., Wheatley, I., and Eperon, I. (1990). Hierarchy for 5 splice site preference determined in vivo. *Journal of molecular biology*, 211(1):103–115.
- Lewis, J. D., Izaurralde, E., Jarmolowski, A., McGuigan, C., and Mattaj, I. W. (1996). A nuclear cap-binding complex facilitates association of u1 snrnp

## BIBLIOGRAPHY

---

- with the cap-proximal 5'splice site. *Genes & Development*, 10(13):1683–1698.
- Li, C. (2016). Association of u2af35, u2af65 and u2 snrnp with constitutively and alternatively spliced pre-mrna. *Doctor of Philosophy*.
- Li, Y. and Blencowe, B. J. (1999). Distinct factor requirements for exonic splicing enhancer function and binding of u2af to the polypyrimidine tract. *Journal of Biological Chemistry*, 274(49):35074–35079.
- Lichtman, J. W. and Conchello, J.-A. (2005). Fluorescence microscopy. *Nature methods*, 2(12):910–919.
- Lim, L. P. and Burge, C. B. (2001). A computational analysis of sequence features involved in recognition of short introns. *Proceedings of the National Academy of Sciences*, 98(20):11193–11198.
- Lin, S. and Fu, X.-D. (2006). Sr proteins and related factors in alternative splicing. *Advances in experimental medicine and biology*, 623:107–122.
- Liu, Z., Luyten, I., Bottomley, M. J., Messias, A. C., Hounginou-Molango, S., Sprangers, R., Zanier, K., Krämer, A., and Sattler, M. (2001). Structural basis for recognition of the intron branch site rna by splicing factor 1. *Science*, 294(5544):1098–1102.
- Long, J. C. and Cáceres, J. F. (2009). The sr protein family of splicing factors: master regulators of gene expression. *Biochemical Journal*, 417(1):15–27.
- Longman, D., McGarvey, T., McCracken, S., Johnstone, I. L., Blencowe, B. J., and Cáceres, J. F. (2001). Multiple interactions between srm160 and sr fam-

## BIBLIOGRAPHY

---

- ily proteins in enhancer-dependent splicing and development of *c. elegans*. *Current Biology*, 11(24):1923–1933.
- MacMillan, A. M., Query, C. C., Allerson, C. R., Chen, S., Verdine, G. L., and Sharp, P. A. (1994). Dynamic association of proteins with the pre-mrna branch region. *Genes & Development*, 8(24):3008–3020.
- Mamatis, T. and Reed, R. (1987). The role of small nuclear ribonucleoprotein particles in pre-mrna splicing. *Nature*, 325:673.
- Maniatis, T. and Reed, R. (2002). An extensive network of coupling among gene expression machines. *Nature*, 416(6880):499–506.
- Manley, J. L. and Krainer, A. R. (2010). A rational nomenclature for serine/arginine-rich protein splicing factors (sr proteins). *Genes & development*, 24(11):1073–1074.
- Manley, J. L., Tacke, R., Hogan, B. L., Friedman, K. L., Diller, J. D., Ferguson, B. M., Nyland, S. V., Brewer, B. J., Fangman, W. L., Saavedra, C., et al. (1996). Sr proteins and splicing control 1569. *Genes & development*.
- Matera, A. G. and Wang, Z. (2014). A day in the life of the spliceosome. *Nature reviews Molecular cell biology*, 15(2):108–121.
- Matlin, A. J., Clark, F., and Smith, C. W. (2005). Understanding alternative splicing: towards a cellular code. *Nature reviews Molecular cell biology*, 6(5):386–398.
- Matlin, A. J. and Moore, M. J. (2006). Spliceosome assembly and composition. *Advances in experimental medicine and biology*, 623:14–35.

## BIBLIOGRAPHY

---

- 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):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):365–375.
- Mayeda, A., Munroe, S., Caceres, J., and Krainer, A. (1994). Function of conserved domains of hnrnp a1 and other hnrnp a/b proteins. *The EMBO journal*, 13(22):5483.
- Mayeda, A., Screatton, G. R., Chandler, S. D., Fu, X.-D., and Krainer, A. R. (1999). Substrate specificities of sr proteins in constitutive splicing are determined by their rna recognition motifs and composite pre-mrna exonic elements. *Molecular and cellular Biology*, 19(3):1853–1863.
- Mayeda, A., Zahler, A. M., Krainer, A. R., and Roth, M. B. (1992). Two members of a conserved family of nuclear phosphoproteins are involved in pre-mrna splicing. *Proceedings of the National Academy of Sciences*, 89(4):1301–1304.
- McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., Bentley, D. L., et al. (1997a). 5-capping enzymes are targeted to pre-mrna by binding to the phosphorylated carboxy-terminal domain of rna polymerase ii. *Genes & development*, 11(24):3306–3318.
- McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S. D., Wickens, M., and Bentley, D. L. (1997b). The c-terminal

## BIBLIOGRAPHY

---

- domain of rna polymerase ii couples mrna processing to transcription. *Nature*, 385(6614):357–361.
- McCracken, S., Lambermon, M., and Blencowe, B. J. (2002). Srm160 splicing coactivator promotes transcript 3-end cleavage. *Molecular and cellular biology*, 22(1):148–160.
- Melhuish, W. (1984). Nomenclature, symbols, units and their usage in spectrochemical analysis-part vi: molecular luminescence spectroscopy. *Pure and Applied Chemistry*, 56(2):231–245.
- 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):5263–5269.
- Mermoud, J. E., Cohen, P., and Lamond, A. I. (1994). Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. *The EMBO journal*, 13(23):5679.
- Michaud, S. and Reed, R. (1991). An atp-independent complex commits pre-mrna to the mammalian spliceosome assembly pathway. *Genes & development*, 5(12b):2534–2546.
- 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):1008–1020.
- Miller, D. M. and Shakes, D. C. (1995). Immunofluorescence microscopy. *Methods in cell biology*, 48:365–394.

## BIBLIOGRAPHY

---

- Mizutani, S. and Temin, H. M. (1970). An rna-dependent dna polymerase in virions of rous sarcoma virus. In *Cold Spring Harbor Symposia on Quantitative Biology*, volume 35, pages 847–849. Cold Spring Harbor Laboratory Press.
- Moerner, W. (2002). A dozen years of single-molecule spectroscopy in physics, chemistry, and biophysics. *The Journal of Physical Chemistry B*, 106(5):910–927.
- Moerner, W. and Fromm, D. P. (2003). Methods of single-molecule fluorescence spectroscopy and microscopy. *Review of Scientific Instruments*, 74(8):3597–3619.
- Monici, M. (2005). Cell and tissue autofluorescence research and diagnostic applications. *Biotechnology annual review*, 11:227–256.
- Moore, M. J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. *Science*, 309(5740):1514–1518.
- 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):364–368.
- Mount, S. M. (1982). A catalogue of splice junction sequences. *Nucleic acids research*, 10(2):459–472.
- Mueller, W. F. and Hertel, K. J. (2011). The role of sr and sr-related proteins in pre-mRNA splicing. In *RNA Binding Proteins*, pages 27–46.
- Mujumdar, R. B., Ernst, L. A., Mujumdar, S. R., Lewis, C. J., and Waggoner,

## BIBLIOGRAPHY

---

- A. S. (1993). Cyanine dye labeling reagents: sulfoindocyanine succinimidyl esters. *Bioconjugate chemistry*, 4(2):105–111.
- Mullen, M., Smith, C., Patton, J., and Nadal-Ginard, B. (1991). Alpha-tropomyosin mutually exclusive exon selection: competition between branchpoint/polypyrimidine tracts determines default exon choice. *Genes & development*, 5(4):642–655.
- Myong, S., Bruno, M. M., Pyle, A. M., and Ha, T. (2007). Spring-loaded mechanism of dna unwinding by hepatitis c virus ns3 helicase. *Science*, 317(5837):513–516.
- 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*, 87(16):6253–6257.
- Nilsen, T. W. (2003). The spliceosome: the most complex macromolecular machine in the cell? *Bioessays*, 25(12):1147–1149.
- Nogués, G., Muñoz, M. J., and Kornblihtt, A. R. (2003). Influence of polymerase ii processivity on alternative splicing depends on splice site strength. *Journal of Biological Chemistry*, 278(52):52166–52171.
- Okunola, H. L. and Krainer, A. R. (2009). Cooperative-binding and splicing-repressive properties of hnrnp a1. *Molecular and cellular biology*, 29(20):5620–5631.
- Olofsson, L. and Margeat, E. (2013). Pulsed interleaved excitation fluorescence spectroscopy with a supercontinuum source. *Optics express*, 21(3):3370–3378.

## BIBLIOGRAPHY

---

- Orphanides, G. and Reinberg, D. (2002). A unified theory of gene expression. *Cell*, 108(4):439–451.
- OMullane, 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):7510–7520.
- Padgett, R. A. (2012). New connections between splicing and human disease. *Trends in Genetics*, 28(4):147–154.
- 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):101–107.
- Pagani, F. and Baralle, F. E. (2004). Genomic variants in exons and introns: identifying the splicing spoilers. *Nature Reviews Genetics*, 5(5):389–396.
- Paule, M. R. and White, R. J. (2000). Survey and summary transcription by rna polymerases i and iii. *Nucleic acids research*, 28(6):1283–1298.
- Pérez-Ortín, J. E., Alepuz, P., Chávez, S., and Choder, M. (2013). Eukaryotic mrna decay: methodologies, pathways, and links to other stages of gene expression. *Journal of molecular biology*, 425(20):3750–3775.
- Plass, M. and Eyraas, E. (2014). Approaches to link rna secondary structures with splicing regulation. *Spliceosomal Pre-mRNA Splicing: Methods and Protocols*, pages 341–356.
- Porrua, O. and Libri, D. (2013). Rna quality control in the nucleus: the angels’ share of rna. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1829(6):604–611.

## BIBLIOGRAPHY

---

- Proudfoot, N. J., Furger, A., and Dye, M. J. (2002). Integrating mrna processing with transcription. *Cell*, 108(4):501–512.
- 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):587–597.
- Ramchatesingh, J., Zahler, A. M., Neugebauer, K. M., Roth, M. B., and Cooper, T. A. (1995). A subset of sr proteins activates splicing of the cardiac troponin t alternative exon by direct interactions with an exonic enhancer. *Molecular and Cellular Biology*, 15(9):4898–4907.
- Reed, R. (1989). The organization of 3' splice-site sequences in mammalian introns. *Genes & development*, 3(12b):2113–2123.
- Reed, R. (1990). Protein composition of mammalian spliceosomes assembled in vitro. *Proceedings of the National Academy of Sciences*, 87(20):8031–8035.
- Reed, R. and Maniatis, T. (1986). A role for exon sequences and splice-site proximity in splice-site selection. *Cell*, 46(5):681–690.
- Reed, R. and Maniatis, T. (1988). The role of the mammalian branchpoint sequence in pre-mrna splicing. *Genes & development*, 2(10):1268–1276.
- Rendell, D. and Mowthorpe, D. J. (1987). *Fluorescence and phosphorescence spectroscopy*. Published on behalf of ACOL, London by Wiley.
- Resch-Genger, U., Grabolle, M., Cavaliere-Jaricot, S., Nitschke, R., and Nann, T. (2008). Quantum dots versus organic dyes as fluorescent labels. *Nature methods*, 5(9):763–775.

## BIBLIOGRAPHY

---

- Rhode, B. M., Hartmuth, K., Westhof, E., and Lührmann, R. (2006). Proximity of conserved u6 and u2 snrna elements to the 5 splice site region in activated spliceosomes. *The EMBO Journal*, 25(11):2475–2486.
- Rino, J., Carvalho, T., Braga, J., Desterro, J. M., Lührmann, R., and Carmo-Fonseca, M. (2007). A stochastic view of spliceosome assembly and recycling in the nucleus. *PLoS Comput Biol*, 3(10):e201.
- 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):84–94.
- Roca, X., Akerman, M., Gaus, H., Berdeja, A., Bennett, C. F., and Krainer, A. R. (2012). Widespread recognition of 5 splice sites by noncanonical base-pairing to u1 snrna involving bulged nucleotides. *Genes & development*, 26(10):1098–1109.
- 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):176–182.
- Roca, X., Krainer, A. R., and Eperon, I. C. (2013). Pick one, but be quick: 5 splice sites and the problems of too many choices. *Genes & development*, 27(2):129–144.
- Rogers, J. and Wall, R. (1980). A mechanism for rna splicing. *Proceedings of the National Academy of Sciences*, 77(4):1877–1879.
- Roscigno, R., Weiner, M., and Garcia-Blanco, M. A. (1993). A mutational analysis of the polypyrimidine tract of introns. effects of sequence differ-

## BIBLIOGRAPHY

---

- ences in pyrimidine tracts on splicing. *Journal of Biological Chemistry*, 268(15):11222–11229.
- Roy, R., Hohng, S., and Ha, T. (2008). A practical guide to single-molecule fret. *Nature methods*, 5(6):507–516.
- 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):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):207–219.
- Rutz, B. and Séraphin, B. (1999). Transient interaction of bbp/scsf1 and mud2 with the splicing machinery affects the kinetics of spliceosome assembly. *Rna*, 5(06):819–831.
- Sako, Y., Minoghchi, S., and Yanagida, T. (2000). Single-molecule imaging of egfr signalling on the surface of living cells. *Nature cell biology*, 2(3):168–172.
- Saltzman, A. L., Pan, Q., and Blencowe, B. J. (2011). Regulation of alternative splicing by the core spliceosomal machinery. *Genes & development*, 25(4):373–384.
- Sanford, J. R., Gray, N. K., Beckmann, K., and Cáceres, J. F. (2004). A novel role for shuttling sr proteins in mrna translation. *Genes & development*, 18(7):755–768.
- Saulière, J., Sureau, A., Expert-Bezançon, A., and Marie, J. (2006). The polypyrimidine tract binding protein (ptb) represses splicing of exon 6b

## BIBLIOGRAPHY

---

- from the  $\beta$ -tropomyosin pre-mrna by directly interfering with the binding of the u2af65 subunit. *Molecular and cellular biology*, 26(23):8755–8769.
- Schaal, T. D. and Maniatis, T. (1999a). Multiple distinct splicing enhancers in the protein-coding sequences of a constitutively spliced pre-mrna. *Molecular and cellular biology*, 19(1):261–273.
- Schaal, T. D. and Maniatis, T. (1999b). Selection and characterization of pre-mrna splicing enhancers: identification of novel sr protein-specific enhancer sequences. *Molecular and cellular biology*, 19(3):1705–1719.
- Schellenberg, M. J., Ritchie, D. B., and MacMillan, A. M. (2008). Pre-mrna splicing: a complex picture in higher definition. *Trends in biochemical sciences*, 33(6):243–246.
- Schenk, G. H. (1973). Absorption of light and ultraviolet radiation fluorescence and phosphorescence emission.
- Schmidt, T., Schütz, G., Baumgartner, W., Gruber, H., and Schindler, H. (1996). Imaging of single molecule diffusion. *Proceedings of the National Academy of Sciences*, 93(7):2926–2929.
- Schoenberg, D. R. and Maquat, L. E. (2012). Regulation of cytoplasmic mrna decay. *Nature Reviews Genetics*, 13(4):246–259.
- Schuler, B. and Eaton, W. A. (2008). Protein folding studied by single-molecule fret. *Current opinion in structural biology*, 18(1):16–26.
- Scott, A., Petrykowska, H. M., Hefferon, T., Gotea, V., and Elnitski, L. (2012). Functional analysis of synonymous substitutions predicted to affect splicing of the cftr gene. *Journal of Cystic Fibrosis*, 11(6):511–517.

## BIBLIOGRAPHY

---

- S raphin, B., Kretzner, L., and Rosbash, M. (1988). A u1 snrna: pre-mrna base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5'cleavage site. *The EMBO Journal*, 7(8):2533.
- Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. *Nature methods*, 2(12):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):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):183–191.
- Sharma, S., Maris, C., Allain, F. H.-T., and Black, D. L. (2011). U1 snrna directly interacts with polypyrimidine tract-binding protein during splicing repression. *Molecular cell*, 41(5):579–588.
- 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):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):367–376.
- Shera, E. B., Seitzinger, N. K., Davis, L. M., Keller, R. A., and Soper, S. A.

## BIBLIOGRAPHY

---

- (1990). Detection of single fluorescent molecules. *Chemical Physics Letters*, 174(6):553–557.
- Sheth, N., Roca, X., Hastings, M. L., Roeder, T., Krainer, A. R., and Sachidanandam, R. (2006). Comprehensive splice-site analysis using comparative genomics. *Nucleic acids research*, 34(14):3955–3967.
- Siliciano, P. G. and Guthrie, C. (1988). 5' splice site selection in yeast: genetic alterations in base-pairing with u1 reveal additional requirements. *Genes & Development*, 2(10):1258–1267.
- Smith, C. W. and Valcárcel, J. (2000). Alternative pre-mrna splicing: the logic of combinatorial control. *Trends in biochemical sciences*, 25(8):381–388.
- Soller, M. (2006). Pre-messenger rna processing and its regulation: a genomic perspective. *Cellular and Molecular Life Sciences CMLS*, 63(7-8):796–819.
- Spena, S., Tenchini, M. L., and Buratti, E. (2006). Cryptic splice site usage in exon 7 of the human fibrinogen b $\beta$ -chain gene is regulated by a naturally silent sf2/asf binding site within this exon. *Rna*, 12(6):948–958.
- 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):55–64.
- Stamm, S. (2008). Regulation of alternative splicing by reversible protein phosphorylation. *Journal of Biological Chemistry*, 283(3):1223–1227.
- Steitz, J. A., Dreyfuss, G., Krainer, A. R., Lamond, A. I., Matera, A. G., and Padgett, R. A. (2008). Where in the cell is the minor spliceosome? *Proceedings of the National Academy of Sciences*, 105(25):8485–8486.

## BIBLIOGRAPHY

---

- Stokes, G. G. (1852). On the change of refrangibility of light. *Philosophical Transactions of the Royal Society of London*, 142:463–562.
- 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):2598–2608.
- Szymczynna, B. R., Bowman, J., McCracken, S., Pineda-Lucena, A., Lu, Y., Cox, B., Lambermon, M., Graveley, B. R., Arrowsmith, C. H., and Blencowe, B. J. (2003). Structure and function of the pwi motif: a novel nucleic acid-binding domain that facilitates pre-mrna processing. *Genes & development*, 17(4):461–475.
- Tacke, R., Tohyama, M., Ogawa, S., and Manley, J. L. (1998). Human tra2 proteins are sequence-specific activators of pre-mrna splicing. *Cell*, 93(1):139–148.
- Tange, T. Ø., Nott, A., and Moore, M. J. (2004). The ever-increasing complexities of the exon junction complex. *Current opinion in cell biology*, 16(3):279–284.
- 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):2704–2717.
- Tazi, J., Bakkour, N., and Stamm, S. (2009). Alternative splicing and disease. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1792(1):14–26.
- Teale, F. and Weber, G. (1957). Ultraviolet fluorescence of the aromatic amino acids. *Biochemical Journal*, 65(3):476.

## BIBLIOGRAPHY

---

- Temin, H. M. and Mizutani, S. (2010). Rna-dependent dna polymerase in virions of rous sarcoma virus. *A century of nature: twenty-one discoveries that changed science and the world*, 181.
- Thompson, N. L. and Steele, B. L. (2007). Total internal reflection with fluorescence correlation spectroscopy. *Nature protocols*, 2(4):878–890.
- Thomson, N. L., Pearce, K. H., and Hsieh, H. V. (1993). Total internal reflection fluorescence microscopy: application to substrate-supported planar membranes. *European biophysics journal*, 22(5):367–378.
- Tian, M. and Maniatis, T. (1993). A splicing enhancer complex controls alternative splicing of doublesex pre-mrna. *Cell*, 74(1):105–114.
- Tilgner, H., Knowles, D. G., Johnson, R., Davis, C. A., Chakraborty, S., Djebali, S., Curado, J., Snyder, M., Gingeras, T. R., and Guigó, R. (2012). Deep sequencing of subcellular rna fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncrnas. *Genome research*, 22(9):1616–1625.
- Tsai, A., Petrov, A., Marshall, R. A., Korlach, J., Uemura, S., and Puglisi, J. D. (2012). Heterogeneous pathways and timing of factor departure during translation initiation. *Nature*, 487(7407):390–393.
- Umen, J. G. and Guthrie, C. (1995). The second catalytic step of pre-mrna splicing. *Rna*, 1(9):869.
- Valcárcel, J., Gaur, R. K., Singh, R., and Green, M. R. (1996). Interaction of u2af65 rs region with pre-mrna of branch point and promotion base pairing with u2 snrna. *Science*, 273(5282):1706.

## BIBLIOGRAPHY

---

- Valcárcel, J. and Green, M. R. (1996). The sr protein family: pleiotropic functions in pre-mrna splicing. *Trends in biochemical sciences*, 21(8):296–301.
- Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., and Luscombe, N. M. (2009). A census of human transcription factors: function, expression and evolution. *Nature Reviews Genetics*, 10(4):252–263.
- Wagner, E. J. and Garcia-Blanco, M. A. (2001). Polypyrimidine tract binding protein antagonizes exon definition. *Molecular and cellular biology*, 21(10):3281–3288.
- Wahl, M. C., Will, C. L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic rnp machine. *Cell*, 136(4):701–718.
- Wang, H.-Y., Lin, W., Dyck, J. A., Yeakley, J. M., Songyang, Z., Cantley, L. C., and Fu, X.-D. (1998). Srp2: a differentially expressed sr protein-specific kinase involved in mediating the interaction and localization of pre-mrna splicing factors in mammalian cells. *The Journal of cell biology*, 140(4):737–750.
- Wang, J. and Manley, J. L. (1995). Overexpression of the sr proteins asf/sf2 and sc35 influences alternative splicing in vivo in diverse ways. *Rna*, 1(3):335–346.
- Wang, Z., Hoffmann, H., and Grabowski, P. (1995). Intrinsic u2af binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *Rna*, 1(1):21–35.
- Wang, Z., Xiao, X., Van Nostrand, E., and Burge, C. B. (2006). General and

## BIBLIOGRAPHY

---

- specific functions of exonic splicing silencers in splicing control. *Molecular cell*, 23(1):61–70.
- Wassarman, D. A. and Steitz, J. A. (1992). Interactions of small nuclear rna's with precursor messenger rna during in vitro splicing. *Science*, 257(5078):1918–1925.
- Watson, J. and Crick, F. (1958). On protein synthesis. In *The Symposia of the Society for Experimental Biology*, volume 12, pages 138–163.
- Weber, G. and Teale, F. (1957). Determination of the absolute quantum yield of fluorescent solutions. *Transactions of the Faraday Society*, 53:646–655.
- Weinmeister, R. (2014). Development of single molecule methods for rna splicing. *Doctor of Philosophy*.
- Widengren, J. and Rigler, R. (1996). Mechanisms of photobleaching investigated by fluorescence correlation spectroscopy. *Bioimaging*, 4(3):149–157.
- Wiederschain, G. Y. (2011). The molecular probes handbook. a guide to fluorescent probes and labeling technologies. *Biochemistry (Moscow)*, 76(11):1276–1276.
- Will, C. L. and Lührmann, R. (2011). Spliceosome structure and function. *Cold Spring Harbor perspectives in biology*, 3(7):a003707.
- Woychik, N. A. and Hampsey, M. (2002). The rna polymerase ii machinery: structure illuminates function. *Cell*, 108(4):453–463.
- Wu, J. and Manley, J. L. (1989). Mammalian pre-mrna branch site selection by u2 snrnp involves base pairing. *Genes & Development*, 3(10):1553–1561.

## BIBLIOGRAPHY

---

- Wu, J. Y. and Maniatis, T. (1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*, 75(6):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):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):334–344.
- Xiao, S.-H. and Manley, J. L. (1998). Phosphorylation–dephosphorylation differentially affects activities of splicing factor ASF/SF2. *The EMBO Journal*, 17(21):6359–6367.
- Yeakley, J. M., Morfin, J.-P., Rosenfeld, M. G., and Fu, X.-D. (1996). A complex of nuclear proteins mediates SR protein binding to a purine-rich splicing enhancer. *Proceedings of the National Academy of Sciences*, 93(15):7582–7587.
- Yildiz, A., Forkey, J. N., McKinney, S. A., Ha, T., Goldman, Y. E., and Selvin, P. R. (2003). Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science*, 300(5628):2061–2065.
- Zahler, A. M. and Roth, M. B. (1995). Distinct functions of SR proteins in recruitment of U1 small nuclear ribonucleoprotein to alternative 5' splice sites. *Proceedings of the National Academy of Sciences*, 92(7):2642–2646.

## BIBLIOGRAPHY

---

- 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*, 86(23):9243–9247.
- Zamore, P. D. and Green, M. R. (1991). Biochemical characterization of u2 snrnp auxiliary factor: an essential pre-mrna splicing factor with a novel intranuclear distribution. *The EMBO Journal*, 10(1):207.
- Zamore, P. D., Patton, J. G., and Green, M. R. (1992). Cloning and domain structure of the mammalian splicing factor u2af. *Nature*, 355:609–614.
- Zhang, B., Kirov, S., and Snoddy, J. (2005a). Webgestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic acids research*, 33(suppl 2):W741–W748.
- Zhang, D. and Rosbash, M. (1999). Identification of eight proteins that cross-link to pre-mrna in the yeast commitment complex. *Genes & Development*, 13(5):581–592.
- Zhang, J. and Powell, S. N. (2005). The role of the brca1 tumor suppressor in dna double-strand break repair. *Molecular Cancer Research*, 3(10):531–539.
- Zhang, X. H., Leslie, C. S., and Chasin, L. A. (2005b). Dichotomous splicing signals in exon flanks. *Genome Research*, 15(6):768–779.
- Zhang, Y., Lu, H., and Bargmann, C. I. (2005c). Pathogenic bacteria induce aversive olfactory learning in caenorhabditis elegans. *Nature*, 438(7065):179–184.
- Zhang, Z., Niu, Z., Yuan, W., Zhao, J., Jiang, F., Zhang, J., Chai, B., Cui, F., Chen, W., Lian, C., et al. (2005d). A mutation in sart3 gene in a chinese

## BIBLIOGRAPHY

---

- pedigree with disseminated superficial actinic porokeratosis. *British Journal of Dermatology*, 152(4):658–663.
- Zhao, J., Hyman, L., and Moore, C. (1999). Formation of mrna 3 ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mrna synthesis. *Microbiology and Molecular Biology Reviews*, 63(2):405–445.
- Zheng, X., Cho, S., Moon, H., Loh, T. J., Oh, H. K., Green, M. R., and Shen, H. (2014). Polypyrimidine tract binding protein inhibits igm pre-mrna splicing by diverting u2 snrna base-pairing away from the branch point. *RNA*, 20(4):440–446.
- Zhong, X.-Y., Wang, P., Han, J., Rosenfeld, M. G., and Fu, X.-D. (2009). Sr proteins in vertical integration of gene expression from transcription to rna processing to translation. *Molecular cell*, 35(1):1–10.
- Zhu, J. and Krainer, A. R. (2000). Pre-mrna splicing in the absence of an sr protein rs domain. *Genes & development*, 14(24):3166–3178.
- Zhu, J., Mayeda, A., and Krainer, A. R. (2001). Exon identity established through differential antagonism between exonic splicing silencer-bound hn-rnp a1 and enhancer-bound sr proteins. *Molecular cell*, 8(6):1351–1361.
- Zhuang, Y. and Weiner, A. M. (1986). A compensatory base change in u1 snrna suppresses a 5 splice site mutation. *Cell*, 46(6):827–835.
- 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):1356–1368.

## BIBLIOGRAPHY

---

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*, 91(8):3363–3367.