

**Mechanism of Action of GGA, a Targeted Oligonucleotide Enhancer of
Splicing Developed for the Treatment of Spinal Muscular Atrophy**

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

Rachel Leanna Dickinson BSc (Hons)
Department of Biochemistry
University of Leicester

July 2013

ABSTRACT

Mechanism of Action of GGA, a Targeted Oligonucleotide Enhancer of Splicing

Developed for the Treatment of Spinal Muscular Atrophy – Rachel Leeanna Dickinson

Spinal muscular atrophy is the leading genetic cause of infant death, and much research has gone into the development of potential therapies for the disease. It is caused by a loss of the SMN1 gene. However, patients have the SMN2 gene, which contains a few silent mutations causing the skipping of exon 7 during splicing. One of the most promising therapeutic strategies involves the use of antisense oligonucleotides to rescue the splicing of SMN2 exon 7, allowing for production of full length SMN protein. One successful antisense oligonucleotide strategy involves the use of a bifunctional targeted oligonucleotide enhancer of splicing (TOES) (Skordis *et al.*, 2003). This oligonucleotide, named GGA, consists of an annealing region that targets it to SMN2 exon 7 and a non-annealed enhancer tail domain, designed to recruit activator proteins and stimulate inclusion of exon 7. However, the precise mechanism by which GGA induces exon 7 inclusion was not fully understood at the time of design. This study has focused on investigation of the mechanism of action of GGA, in order to improve the therapeutic potential for GGA and future TOES. GGA was found to bind directly to SRSF1, an activator protein, via its enhancer tail domain. The tail domain forms a G-quadruplex structure *in vitro* (Smith *et al.*, manuscript submitted). The presence of this structure in nuclear extracts was confirmed, and the enhancer domain was found to bind the G-quadruplex associated proteins CNBP and nucleolin. Stabilization of this structure using ligands reduced the efficacy of GGA, indicating that GGA does not form a G-quadruplex when it is actively stimulating SMN2 exon 7 inclusion. Single molecule methods revealed that the annealing domain of GGA, which anneals over an exonic splicing silencer shown to bind hnRNP A1 and/or Sam68 (Kashima *et al.*, 2007; Pedrotti *et al.*, 2010), reduces the number of Sam68 proteins bound per SMN2 RNA. These findings are consistent with the fact that the annealing region of GGA promotes U2AF65 binding and the enhancer tail domain promotes U2 snRNP binding to SMN2 transcripts (Smith *et al.*, manuscript submitted; Smith, 2012).

ACKNOWLEDGEMENTS

I consider myself extremely fortunate and grateful to have had not one but two wonderful supervisors who were always there to provide support and guidance when it was needed. Dr Glenn Burley, my chemistry supervisor, played a very significant role in the early days of my studies and, even after moving to the University of Strathclyde, made sure to keep in touch. Professor Ian Eperon, my biochemistry supervisor, was always available and more than happy to devote as much time as necessary to listen to and help me with whatever problems I was having, both professional and personal, throughout my PhD. I would like to thank Peter Moody, my third committee member, for his helpful discussions and insights during meetings.

I would also like to thank my collaborators for all of their contributions. Dr Cyril Dominguez, in addition to providing hnRNP F samples, also provided many helpful suggestions and useful discussions. I am grateful to Dr Haiyan Zhou for collaborating with me and for providing various SMN2-targeting antisense oligonucleotides. I am also grateful to Professor Laurence Hurley for providing G-quadruplex stabilizing ligands. The collaborations of Professor Mark Searle and his student Alex Cousins, which were vital to the G-quadruplex investigations of GGA, are greatly appreciated. I also appreciate all of the mass spectrometry help provided by Shairbanu Ibrahim.

I would like to give special thanks to the many lab members, both in organic chemistry and biochemistry, who helped to create a fun, friendly, and productive atmosphere during my PhD. I am especially grateful to Andrew Perrett for all of his guidance, patience and friendship during my time in the chemistry department and afterwards. I could never thank Dr Lindsay Smith enough for her support, influence and friendship both in and out of the lab. I am also grateful to Dr Lucy Eperon, Christian Lucas, Robert Weinmeister, Dr Mark Hodson, Paul Smith and Sheree Purmessur for their friendship, practical support and helpful discussions.

Last but not least, I would like to thank my husband Warwick Dumas, my parents Lynda and Britt Dickinson, my sister Amanda Dickinson and all of the rest of my family for being supportive and loving throughout my PhD.

ABBREVIATIONS

ABBREVIATIONS

AMPS – Ammonium persulfate

ASO – Antisense oligonucleotide

ATP – Adenosine-5'-triphosphate

BABE – Bromoacetamidobenzyl-EDTA

bp – Base pair

BPS – Branch point sequence

BSA – Bovine serum albumin

CD – Circular dichroism

CNBP – Cellular nucleic acid-binding protein

CNS – Central nervous system

CrPi – Creatine phosphate

Cy5 – Indocarbocyanin[®]

ddNTPs – Dideoxyribonucleotide triphosphates

DI – Deionized

DMEM – Dulbecco's modified Eagle's medium

DMSO – Dimethyl sulfoxide

DNA – Deoxyribonucleic acid

dNTPs – Deoxyribonucleotide triphosphates

DTT – Dithiothreitol

ECL – Enhanced chemiluminescence

EColi – Escherichia coli

EDTA – Ethylenediamine-N,N,N',N'-tetraacetic acid

eGFP – Enhanced green fluorescent protein

EMCCD – Electron-multiplying charge-coupled device

ESE – Exonic splicing enhancer

ABBREVIATIONS

ESS – Exonic splicing silencer

EtOH – Ethanol

Fe-BABE – Iron chelate of bromoacetamidobenzyl-EDTA

FBS – Fetal bovine serum

FMRP – Fragile X mental retardation protein

GFP – Green fluorescent protein

GNP – Gold nanoparticle

GTP – Guanosine-5'-triphosphate

h – Hour

HEG – Hexaethylene glycol

HEK 293T – Human embryonic kidney 293T

Hepes – N-2-hydroxyethylpiperazine-N-2-ethane sulfonic acid

hnRNP – Heterogeneous nuclear ribonucleoprotein

HPLC – High-performance liquid chromatography

HRP – Horseradish peroxidase

ISE – Intronic splicing enhancer

ISS – Intronic splicing silencer

kDa – Kilodalton

K/glu – Potassium glutamate

l – Litre

LC MS/MS – Liquid chromatography - tandem mass spectrometry

LMP – Low melting point

LNA – Locked nucleic acid

ABBREVIATIONS

m – Milli

M – Molar

MALDI-TOF – Matrix-assisted laser desorption/ionization – time of flight

min(s) – Minute(s)

mRNA – Messenger RNA

n – Nano

NCL – Nucleolin

NE – Nuclear extract

NMD – Nonsense-mediated Decay

NMR – Nuclear magnetic resonance

NP-40 – Nonidet P-40

nt – Nucleotide

Oligo – Oligonucleotide

ON – Overnight

2'OMe – 2' O-Methyl

PAGE – Polyacrylamide gel electrophoresis

PBS – Phosphate buffered saline

PCA – Protocatechuic acid

PCD – Protocatechuate dioxygenase

PCR – Polymerase chain reaction

PCV – Packed cell volume

PEG – Poly(ethylene glycol)

PEI – Poly(ethylene imine)

PK – Proteinase K

PMO – Phosphoramidate morpholino

PNACL – Protein Nucleic Acid Chemistry Laboratory

Pre-mRNA – Precursor messenger RNA

PPT – Polypyrimidine tract

ABBREVIATIONS

PTB – Polypyrimidine tract binding protein

RNA – Ribonucleic acid

RNase – Ribonuclease

rNTPs – Ribonucleotide triphosphates

RRM – RNA recognition motif

RT – Room temperature

RTXase – Reverse transcriptase

RTXN – Reverse transcription

RON – Receptor d'origine nantais

Sam68 – Src-associated in mitosis 68 kDa protein

SDS – Sodium dodecyl sulfate

Sec – second(s)

SHAPE – Selective 2'-hydroxyl acylation analyzed by primer extension

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

TBE – Tris-borate-EDTA buffer

TBS – Tris-buffered saline

TCA – Trichloroacetic acid

TCEP – Tris(2-carboxyethyl)phosphine hydrochloride

TEMED – N,N,N',N'-tetramethylethylenediamine

TIRF – Total internal reflection

TLC – Thin layer chromatography

TOES – Tailed oligonucleotide enhancer(s) of splicing

ABBREVIATIONS

U snRNP – Uridine-rich small nuclear ribonucleoprotein particles

U2AF – U2 snRNP auxiliary factor

UV – Ultraviolet

V – Volts

v/v – Volume per volume

W – Watts

w/v – Weight per volume

CONTENTS

ABSTRACT..... i

ACKNOWLEDGEMENTS ii

ABBREVIATIONS..... iii

CONTENTSviii

LIST OF FIGURES.....xiv

1 INTRODUCTION 1

 1.1 Alternative Pre-mRNA Splicing..... 3

 1.1.1 Basic Splicing Machinery and Mechanism..... 5

 1.1.2 Regulation of Splicing 11

 1.2 Alternative Splicing and Disease 23

 1.2.1 Therapeutic Strategies to Influence Alternative Splicing 25

 1.3 Spinal Muscular Atrophy..... 31

 1.3.1 Splicing of SMN1 and SMN2 33

 1.3.2 Therapeutic Approaches..... 34

 1.4 G-Quadruplexes 38

 1.4.1 G-quadruplex Structure 40

 1.5 Specific Aims 43

CONTENTS

2	MATERIALS AND METHODS	44
2.1	<i>In vitro</i> Protein Techniques	45
2.1.1	Biotin Affinity Purification Methods	45
2.1.2	Trichloroacetic Acid Precipitation of Proteins	48
2.1.3	SDS PAGE.....	49
2.1.4	Western Blot Analysis	50
2.1.5	UV Crosslinking with Oligonucleotides	51
2.2	<i>In vitro</i> RNA Techniques	52
2.2.1	Purification Techniques	52
2.2.2	Oligonucleotide Techniques	54
2.2.3	Glycerol Gradients	56
2.2.4	<i>In vitro</i> Transcription.....	58
2.2.5	<i>In vitro</i> Splicing.....	59
2.2.6	Analysis of Splicing Complexes A, B and C by Native Agarose Gel Electrophoresis	63
2.2.7	Quantification of Gels with Radioactive RNA/DNA	63
2.3	Cell Culture and Nuclear Extract Preparation	64
2.3.1	Transfection of HEK 293T Cells	64
2.3.2	Nuclear Extract Preparation	65
2.3.3	Dialysis of Nuclear Extract	65

CONTENTS

2.4	Single Molecule Techniques.....	65
2.4.1	Preparation of the Sample Chamber	65
2.4.2	Single Molecule Experiment	66
2.4.3	Data Acquisition and Analysis.....	67
2.5	Techniques Related to Fe-BABE Experiments.....	69
2.5.1	Synthesis of Fe-BABE Modified Oligonucleotides.....	69
2.5.2	Fe-EDTA experiments	70
2.5.3	Fe-BABE experiments.....	71
3	BIOTIN AFFINITY PURIFICATION WITH PHOTOCLEAVABLE LINKER.....	73
3.1	Introduction	74
3.2	Design of Oligonucleotide with Photocleavable Linker	75
3.3	Optimization of Method	75
3.4	Extension of Method to Purify Intact Complexes	79
3.5	Discussion.....	83
4	MECHANISMS OF ACTION OF A BIFUNCTIONAL OLIGONUCLEOTIDE ENHANCER OF SPLICING	84
4.1	Introduction	85
4.2	Complexes and Bound Proteins	86
4.2.1	Identification of Proteins	86
4.2.2	Binding Specificity and Strength	96

CONTENTS

4.2.3	Protein Binding to GGA is Unchanged in Presence of SMN2 RNA	100
4.3	G-quadruplex Nature of the GGA Tail	106
4.4	Actions of hnRNP F	121
4.5	Importance of Oligonucleotide Chemistry	128
4.6	Analysis of the Bifunctional Nature of GGA	133
4.7	Discussion	138
5	NEW STRATEGIES FOR RESCUING THE SPLICING OF SMN2 EXON 7 USING OLIGONUCLEOTIDES	140
5.1	Introduction	141
5.2	Use of ESE Conjugated Gold Nanoparticles to Enhance SMN2 Exon 7 Inclusion 141	
5.3	<i>In vitro</i> Splicing with Morpholino Oligonucleotides Targeting a Silencer in SMN2 Intron 7	151
5.4	Discussion	155
6	USE OF Fe-BABE TO PROBE SPLICING REACTION MECHANISMS	158
6.1	Introduction	159
6.2	Synthesis of Fe-BABE Conjugated Oligonucleotides	160
6.3	Confirmation of Cleavage Conditions using Fe-EDTA	161
6.4	Fe-BABE Experiments on SMN2 using Fe-BABE Modified Oligonucleotides .	164
6.5	Discussion	170
7	DISCUSSION	171

CONTENTS

7.1	Importance of Understanding the Mechanism by which GGA Enhances SMN2 Exon 7 Inclusion	172
7.2	Stabilization of the G-quadruplex Structure of the GGA Enhancer Tail Domain Decreases Functionality	173
7.3	GGA Binds SRSF1 and G-quadruplex-Associated Proteins and Recruits many Splicing and Transcription-Related Proteins.....	175
7.4	The Enhancer Tail and Annealing Domains of GGA both have Significant Mechanistic Roles	180
7.5	Overview of Proposed Mechanism of Action of GGA.....	182
8	APPENDIX.....	183
8.1	Splicing Template Sequences.....	183
8.1.1	SMN2.....	183
8.1.2	SMN2 Exon 7 'Exon Defined' Construct.....	184
8.1.3	β -globin	184
8.1.4	β -globin + GGA anneal site	185
8.1.5	Adenovirus WT (A2)	185
8.1.6	Adenovirus + ESE (A3).....	186
8.2	Oligonucleotide Sequences.....	186
8.2.1	SMN2 Splice-Altering RNA Oligonucleotides	187
8.2.2	Oligonucleotides for SMN2 Fe-BABE Experiments.....	188
8.2.3	Reverse Transcription Primers.....	188

CONTENTS

8.2.4 Other Oligonucleotides.....	189
BIBLIOGRAPHY	190

LIST OF FIGURES

1 - INTRODUCTION

Figure 1.1: Diagram of the transesterification reactions involved in pre-mRNA splicing	6
Figure 1.2: U1 snRNP as an example of snRNP structure.....	7
Figure 1.3: Spliceosome Assembly	8
Figure 1.4: Key Interactions of Complexes E and A.....	9
Figure 1.5: Diagram of Intron and Exon Definition	20
Figure 1.6: Diagram of the Different Types of Alternative Splicing.....	22
Figure 1.7: Types of Chemical Modifications for Oligonucleotides.....	28
Figure 1.8: Splice-Altering Antisense Oligonucleotide Strategies	30
Figure 1.9: Basic G-quadruplex Structure.....	42

3 - BIOTIN AFFINITY PURIFICATION WITH PHOTOCLEAVABLE LINKER

Figure 3.1: Design of GGA PC Bio, oligonucleotide for biotin affinity purification with photocleavable linker.	75
Figure 3.2: Detergent is vital for clean affinity purifications.....	76
Figure 3.3: Diagram of the assembly used for UV application to affinity purified samples in columns.....	77
Figure 3.4: Application of 365 nm UV light for 10 min is sufficient to cleave the photocleavable linker of GGA PC Bio.....	78

LIST OF FIGURES

Figure 3.5: Affinity purification using a photocleavable linker improves the cleanliness of eluted samples.....	79
Figure 3.6: The gentle photocleavage method of elution using GGA PC Bio allows for elution of an intact complex on the GGA tail.	80
Figure 3.7: Affinity purification of β -globin splicing complex A using GGA PC Bio annealed at the 3' end of β -globin was not successful	82

4 - MECHANISMS OF ACTION OF A BIFUNCTIONAL OLIGONUCLEOTIDE

ENHANCER OF SPLICING

Figure 4.1: List of specific proteins found to be associated with the GGA tail after affinity purification with GGA PC Bio.....	88
Figure 4.2: Comparison of proteins found to be associated with the tail of GGA PC Bio by affinity purification with proteins found to bind directly to GGA and the GGA tail..	89
Figure 4.3: Affinity purification with GGA PC Bio revealed 5'-3' exoribonuclease 2 and nucleolin binding.....	91
Figure 4.4: CNBP binds directly to the enhancer tail of GGA.....	92
Figure 4.5: SRSF1 binds to the tail of GGA	95
Figure 4.6: Affinity purification with GGA PC Bio shows that addition of 50 x GGA successfully competes off all proteins binding to GGA PC Bio	96
Figure 4.7: Analysis of protein binding to GGA-O, O-TO and O-NT by UV crosslinking and native polyacrylamide gel electrophoresis	97
Figure 4.8: Purified SRSF1 can bind 5' radiolabelled GGA-O.....	99

LIST OF FIGURES

Figure 4.9: Detection of purified SRSF1 binding to 5' radiolabelled GGA-O by native polyacrylamide gel electrophoresis is further improved by increasing the amount of glycerol in the gel from 5% (A) to 10% (B).....	100
Figure 4.10: Composition of proteins bound to GGA is unchanged in the presence of SMN2 RNA.....	101
Figure 4.11: Nucleolin binding is greatly reduced when ATP is depleted	102
Figure 4.12: Protein binding to GGA is unchanged in the presence of SMN2 RNA, which has been stalled at splicing complex A; however, GGA does not remain associated with the SMN2 RNA when analyzed by glycerol gradient.	104
Figure 4.13: Annealing region of GGA is very weak.	106
Figure 4.14: Diagram of GGA oligonucleotide with G's capable of forming a G-quadruplex indicated in red.....	107
Figure 4.15: Structures of three G-quadruplex stabilizing ligands. (A) GQC-05 (B) GSA-0820 (C) GSA-0902	108
Figure 4.16: Silanized plasticware is necessary for experiments using G-quadruplex stabilizing ligands to prevent loss of RNA.....	110
Figure 4.17: GQC-05 reduces SMN2 exon 7 inclusion.....	111
Figure 4.18: GSA-0820 reduces SMN2 exon 7 inclusion.	112
Figure 4.19: GSA-0902 reduces SMN2 exon 7 inclusion.	113
Figure 4.20: GSA-0820 and GSA-0902 have a greater inhibitory effect on SMN2 exon 7 inclusion in the presence of GGA, while GQC-05 has a negative effect on SMN2 exon 7 inclusion in the absence of GGA, which is similar or greater than the effect seen in the presence of GGA.	114
Figure 4.21: Protein binding to GGA and TO is unaffected by 2.5 μ M GQC-05.	116

LIST OF FIGURES

Figure 4.22: Protein binding to GGA and TO is unaffected by 2.5 μ M GSA-0820.....	117
Figure 4.23: Protein binding to GGA and TO is unaffected by 2.5 μ M GSA-0902.....	118
Figure 4.24: Secondary structure analysis of the enhancer tail of GGA is consistent with the formation of a G-quadruplex, even in the presence of nuclear extract	120
Figure 4.25: Apparent increase in exon 7 inclusion using purified hnRNP F RRM1 & 2, later shown to be due to problems with control buffer.....	123
Figure 4.26: SMN2 splicing with GGA and purified hnRNP F RRM1 & 2 in Buffer D shows that although all of the components of Buffer D are safe for splicing, the total increase in solutes was still inhibitory	124
Figure 4.27: SMN2 splicing with GGA and hnRNP F RRM1 & 2 in K/glu + Hepes buffer, giving final concentrations as for a normal splicing reaction with the addition of 0.1% β -mercaptoethanol, shows only very slight inhibition of splicing	125
Figure 4.28: SMN2 splicing with GGA and hnRNP F RRM3 in K/glu + Hepes buffer shows very slight negative effect on ratio of exon 7 inclusion/exclusions	126
Figure 4.29: SMN2 splicing with GGA and new batch of hnRNP F RRM1 & 2 in sodium phosphate buffer further proves that the results from Figure 4.25 were false.....	127
Figure 4.30: Oligonucleotides with the same sequence as GGA, but different chemical modifications show significant variations in protein binding.....	130
Figure 4.31: Oligonucleotides with the same sequence as GGA, but different chemical modifications show variations in secondary structure.....	131
Figure 4.32: Loss of phosphorothioate modifications over time in various batches of GGA causes changes in secondary structure and a reduction in the ability to stimulate SMN2 exon 7 inclusion	132
Figure 4.33: GGA PC Bio is capable of enhancing SMN2 exon 7 inclusion.....	134

LIST OF FIGURES

Figure 4.34: Enhancer tail of GGA PC Bio needs to be intact, even after 30 minutes of SMN2 splicing, in order to fully stimulate SMN2 exon 7 inclusion	135
Figure 4.35: The annealing region of GGA is capable of reducing the number of Sam68 molecules bound to a single SMN2 transcript.....	137

5 - NEW STRATEGIES FOR RESCUING THE SPLICING OF SMN2 EXON 7 USING OLIGONUCLEOTIDES

Figure 5.1: Diagram of the gold nanoparticles conjugated with 2'OMe thiol modified oligonucleotides that were used for <i>in vitro</i> SMN2 splicing experiments.....	142
Figure 5.2: Triplicate <i>in vitro</i> SMN2 splicing with increasing concentrations of the gold nanoparticles described in Figure 5.1.....	144
Figure 5.3: Quantification of the SMN2 <i>in vitro</i> splicing results with modified gold nanoparticles from Figure 5.2.....	145
Figure 5.4: Triplicate <i>in vitro</i> A2 splicing with increasing concentrations of GNP 16 (Figure 5.1) and the relevant controls.....	147
Figure 5.5: Quantification of the A2 <i>in vitro</i> splicing results with GNP 16 from Figure 5.4. GNP 16 clearly promotes use of the downstream splice site.....	148
Figure 5.6: Triplicate <i>in vitro</i> A3 splicing with increasing concentrations of GNP 16 (Figure 5.1) and the relevant controls.....	149
Figure 5.7: Quantification of the A3 <i>in vitro</i> splicing results with GNP 16 from Figure 5.6. GNP 16 clearly promotes use of the downstream splice site.....	150

LIST OF FIGURES

Figure 5.8: Triplicate <i>in vitro</i> SMN2 splicing with increasing concentrations of three antisense morpholino oligonucleotides targeting an ISS in exon 7.	153
Figure 5.9: Quantification of the SMN2 <i>in vitro</i> splicing results with PMO18, PMO20, and PMO25 from Figure 5.8. PMO25 performs slightly better than the other two oligonucleotides, especially at 250 nM	154

6 - USE OF Fe-BABE TO PROBE SPLICING REACTION MECHANISMS

Figure 6.1: Conjugation of 5' thiol modified oligonucleotides with Fe-BABE	161
Figure 6.2: Optimization of hydroxyl radical cleavage conditions using Fe-EDTA.	162
Figure 6.3: Glycerol-free dialyzed nuclear extract still allows for formation of GGA-dependent SMN2 exon 7 defined complex.	163
Figure 6.4: Fe-BABE experiment with Fe-BABE GGA-O and Fe-BABE NET-O without nuclear extract shows no specific hydroxyl radical cleavage	165
Figure 6.5: Fe-BABE experiment with Fe-BABE GGA-O and Fe-BABE NET-O with nuclear extract shows no specific hydroxyl radical cleavage	166
Figure 6.6: Optimization of cleavage conditions with Fe-BABE GGA-O and radiolabelled SMN2 RNA, only non-specific cleavage observed.	168
Figure 6.7: Radiolabelled SMN2 cleavage test using Fe-BABE NT and Fe-BABE GGA shows specific hydroxyl radical induced cleavage only with Fe-BABE NT.....	169

1 INTRODUCTION

1.1 Alternative Pre-mRNA Splicing

1.2 Alternative Splicing and Disease

1.3 Spinal Muscular Atrophy

1.4 G-Quadruplexes

1.5 Specific Aims

INTRODUCTION

Spinal muscular atrophy (SMA) is the leading genetic cause of death of infants and involves the loss of the SMN1 gene (Pearn, 1980; Lunn & Wang, 2008; Feldkötter *et al.*, 2002). Due to evolutionary good-fortune, SMA patients have at least one copy of a very similar gene, SMN2, which only differs by a few translationally silent mutations. However, these mutations, especially a C to T transition at position 6 in exon 7 (C6T), cause SMN2 to exclude exon 7 during splicing in 90% of transcripts, leading to the production of a truncated SMN protein that is rapidly degraded (Monani *et al.*, 1999; Lorson *et al.*, 1999; Le *et al.*, 2000; Cartegni *et al.*, 2006; Lorson *et al.*, 1998). This mutation disrupts an exonic splicing enhancer known to bind SRSF1 and creates an exonic splicing silencer, which has been shown to bind hnRNP A1 and/or Sam68 (Lorson *et al.*, 1999; Monani *et al.*, 1999; Cartegni & Krainer, 2002; Cartegni *et al.*, 2006; Kashima & Manley, 2003; Kashima *et al.*, 2007; Pedrotti *et al.*, 2010). Various therapeutic strategies have been developed to restore SMN2 exon 7 inclusion (Section 1.3.2). One strategy involves a bifunctional targeted oligonucleotide enhancer of splicing (TOES) named GGA, which was designed to anneal from position 2-16 on SMN2 exon 7 and recruit activator proteins with an enhancer tail extending from position 16 (AGGAGGACGGAGGACGGAGGACA) (Skordis *et al.*, 2003). This sequence has been successful in stimulating SMN2 exon 7 inclusion both *in vitro* and *in vivo* (Skordis *et al.*, 2003; Meyer *et al.*, 2009). However, the precise mechanism of action of this oligonucleotide was not proven. The purpose of this study was to investigate the exact way in which GGA is able to restore SMN2 exon 7 inclusion, as this knowledge would improve the therapeutic appeal of TOES and provide further insights into how enhancer sequences in general function.

1.1 Alternative Pre-mRNA Splicing

Historically, the view on gene expression was much simpler than today: one gene, one protein. Then in the 1970's RNA splicing was discovered, which led to a whole field of research, ultimately revealing a much more complex system where one gene is capable of producing many different messenger RNAs (mRNAs) and therefore many different protein isoforms (Sharp, 2005; Berget *et al.*, 1977; Chow *et al.*, 1977). Pre-mRNAs contain varying numbers of sequences known as exons and introns. Splicing involves the removal of introns and joining together of exons to produce mRNA for translation. Many pre-mRNA transcripts have the ability to be spliced in multiple different patterns. This process of producing multiple unique mRNAs from one pre-mRNA is called alternative splicing. Levels of alternative splicing and splicing in general vary widely between different organisms, where higher levels of alternative splicing are typically associated with higher levels of organism complexity (Schad *et al.*, 2011). For example, alternative splicing has provided a reassuring explanation for how a simple water flea (*Daphnia pulex*) could have about 31,000 genes (Colbourne *et al.*, 2011), while humans have only around 21,000 (Harrow *et al.*, 2012). Over the years, the estimated number of human genes that are alternatively spliced has continued to increase. Currently it is thought that around 94% of human genes are alternatively spliced, clearly indicating the prevalence and significance of this amazing process (Wang *et al.*, 2008; Pan *et al.*, 2008).

As with any other important biological process, alternative splicing involves a large amount of regulation. However, the full extent and the precise mechanisms of regulation are still not entirely understood (Roca *et al.*, 2013). In addition to many

INTRODUCTION

regulatory proteins, which bind to the pre-mRNA or are associated with proteins that do, there are other factors that influence splicing patterns. For example, as the splicing process begins while the pre-mRNA is still being transcribed, the rate of transcription can affect the pattern of splicing (Eperon *et al.*, 1988; Roberts *et al.*, 1998; Kornblihtt *et al.*, 2004; Dujardin *et al.*, 2013). Therefore, chromatin structure, which can cause pausing of the RNA polymerase, can also influence splicing (Schwartz *et al.*, 2009; Kolasinska-Zwierz *et al.*, 2009; Andersson *et al.*, 2009; Spies *et al.*, 2009). Organisms make the most of this dynamic system, using the regulation of alternative splicing to produce different protein isoforms at different stages of development, in different tissue and cell types, and in response to various external stimuli or to alterations in cellular status (Boutz *et al.*, 2007; Yeo *et al.*, 2004; Stamm, 2002; Shin & Manley, 2004; McGlincy *et al.*, 2012).

1.1.1 Basic Splicing Machinery and Mechanism

Pre-mRNA splicing is carried out by the spliceosome, a complex catalytic machine composed of five core uridine-rich small nuclear ribonucleoproteins (snRNPs) along with around 170 other associated proteins (Wahl *et al.*, 2009). Unlike other cellular machines, such as the ribosome, which are mostly pre-assembled for duty, the spliceosome must be re-assembled on the pre-mRNA for each splicing event. Although spliceosome assembly is a dynamic process, scientists have been able to study its progression using various methods to stall the spliceosome at different stages for analysis.

The spliceosome is responsible for identifying the exons and introns in the pre-mRNA and catalyzing the excision of the intronic sequences via two sequential transesterification reactions (Moore & Sharp, 1993). There are three key sequence elements, which are involved in the transesterification reactions: the 5' splice site (5'ss), characterized by a conserved GU dinucleotide at the 5' exon/intron boundary; the 3' splice site (3'ss), characterized by a conserved AG at the 3' intron/exon boundary; and the branch point sequence found in the intron, characterized by a conserved adenosine (Staley & Guthrie, 1998). In the first transesterification reaction, the 2' OH group of the branch point adenosine performs a nucleophilic attack on phosphate group at the 5'ss, resulting in the formation of an intronic lariat intermediate and release of the 5' exon (Figure 1.1). Subsequently, the newly generated 2' OH group at the 5'ss attacks the 3'ss, resulting in fusion of the two exons and release of the intronic lariat (Figure 1.1) (Staley & Guthrie, 1998; Grabowski *et al.*, 1984; Ruskin *et al.*, 1984; Konarska *et al.*, 1985).

Figure removed for copyright reasons.
Please see hard copy version of this thesis
or reference(s) below for the image.

Figure 1.1: Diagram of the transesterification reactions involved in pre-mRNA splicing. (Staley & Guthrie, 1998)

The first transesterification reaction involves nucleophilic attack of the 2'OH of the branch point adenosine on the phosphodiester bond at the 5'ss, generating the intron lariat. Subsequently, the 5'ss joins the 3'ss in a similar reaction, releasing the lariat.

The spliceosome utilizes five uridine-rich snRNPs (U snRNPs), in addition to many auxiliary factors, to locate and facilitate the two transesterification reactions necessary for each splicing event. These U snRNPs are made up of structured snRNAs bound to a circular core of seven Sm, or Sm-like in the case of U6, proteins as well as various other RNPs associated with each specific U snRNP (Figure 1.2) (Hermann *et al.*, 1995; Raker *et al.*, 1996; Mayes *et al.*, 1999; Kiss, 2004). These U snRNPs do not have pre-formed active sites, and they must work together to identify the correct splicing sequences, often across vast expanses of intronic sequence, and bring them close enough together

INTRODUCTION

to catalyze the transesterification reactions (Hong *et al.*, 2006). This dynamic process involves many factors in addition to the core U snRNPs, which influence splice site selection and aid in the necessary rearrangements of the U snRNPs to create a functional spliceosome (Wahl *et al.*, 2009). Although spliceosome assembly and disassembly in reality should be viewed as a continuous and fluid process, in order to understand the natural progression of this amazing process, several spliceosomal complexes have been elucidated using various techniques to stall spliceosome assembly at different stages (Figure 1.3).

Figure removed for copyright reasons.
Please see hard copy version of this thesis
or reference(s) below for the image.

Figure 1.2: U1 snRNP as an example of snRNP structure

(A) Diagram of U1 snRNP showing the U1 snRNA with its associated proteins, U1A, U1C, and U1 70K as well as the core ring of proteins common to most U snRNPs. **(B)** Ribbon diagram of the core sm proteins common to all of the major U snRNPs, except U6. (Stark *et al.*, 2001; Mura *et al.*, 2001; Zieve & Khusial, 2003)

Figure removed for copyright reasons.
Please see hard copy version of this thesis
or reference(s) below for the image.

Figure 1.3: Spliceosome Assembly (Wahl *et al.*, 2009)

This diagram shows the stepwise progression of spliceosome assembly through various complexes that have been identified experimentally, focusing on the key interactions of the U1 snRNPs.

The early spliceosomal complex, complex E, is characterized by the ATP-independent binding of the U1 snRNP to the 5'ss through base pairing, which is stabilized by the binding of additional factors, such as SR proteins (Reed, 1990; Wassarman & Steitz, 1992; Görnemann *et al.*, 2005; Huranová *et al.*, 2010; Wahl *et al.*, 2009). The 3'ss is also identified in complex E by the binding of U2 auxiliary factor (U2AF), which is composed of two subunits: U2AF65, which recognizes the polypyrimidine tract (PPT) near the 3'ss, and U2AF35, which recognizes the AG dinucleotide of the 3'ss (Figure 1.4) (Kramer & Utans, 1991; Gaur *et al.*, 1995; Rudner *et al.*, 1998; Graveley *et al.*, 2001). Additionally, the branch point sequence (BPS) is bound by SF1, also known as branch point binding protein (BBP) (Kramer & Utans, 1991; Berglund *et al.*, 1998). The

INTRODUCTION

U2 snRNP has been found to be weakly associated through protein-protein interactions in complex E, but it does not become fully engaged, replacing SF1 binding at the BPS, until the progression to the ATP-dependent complex A (Figure 1.4) (Hong *et al.*, 1997; Wahl *et al.*, 2009). This base-pairing interaction of the U2 snRNP to the BPS is stabilized by SF3a and SF3b, which are associated with the U2 snRNP, as well as by the serine-arginine rich domain (RS domain) of U2AF65 (Hastings & Krainer, 2001).

Figure removed for copyright reasons.
Please see hard copy version of this thesis
or reference(s) below for the image.

Figure 1.4: Key Interactions of Complexes E and A (Hastings & Krainer, 2001)

Initially, the 5'ss and 3'ss are recognized by the U1 snRNP and U2AF respectively. Although the U2 snRNP has been found to be associated with some E complexes, it does not base-pair to the branch point sequence until A complex.

INTRODUCTION

Progression to complex B involves the recruitment of the tri-snRNP consisting of U4 and U6 snRNPs, which base pair together, and the U5 snRNP (Bringmann *et al.*, 1984; Hashimoto & Steitz, 1984; Rinke *et al.*, 1985; Wahl *et al.*, 2009). The U6 snRNP then breaks its connection with the U4 snRNP and replaces the U1 snRNP base pairing at the 5'ss, with the help of several auxiliary factors (Figure 1.3) (Maroney *et al.*, 2000). Also, the U6 snRNP forms base pairs with the U2 snRNP, and together with the U5 snRNP forms the catalytically active B* complex (Madhani & Guthrie, 1992; Fortner *et al.*, 1994; Wahl *et al.*, 2009). The first transesterification reaction between the branch point adenosine and the 5'ss can then occur, progressing assembly to complex C. Further rearrangements of the catalytic center formed by the U2, U6 and U5 snRNPs are necessary prior to the final transesterification reaction, which joins the 5' and 3' splice sites and releases the intron lariat (Figure 1.3) (Wahl *et al.*, 2009).

It is worth mentioning that a minor spliceosome also exists, which functions in a similar way to the major spliceosome, but utilizes slightly different U snRNPs and is responsible for splicing a very small set of different introns (Turunen *et al.*, 2013). The U5 snRNP is used in both the major and minor spliceosomes; however, in the minor spliceosome, the U1, U2, U4 and U6 snRNPs are replaced by the U11, U12, U4atac and U6atac snRNPs respectively (Schneider *et al.*, 2002; Will & Lührmann, 2005; Turunen *et al.*, 2013).

As spliceosome assembly is such a dynamic process, which is repeated for every splicing event, it is subject to a large amount of regulation, especially in the early stages of assembly. These key properties of flexibility, complexity and susceptibility to

INTRODUCTION

regulation allow splicing patterns to be altered by a wide range of different circumstances.

1.1.2 Regulation of Splicing

Although pre-mRNAs contain key splicing sequences (5'ss, 3'ss, BPS, and PPT), which are recognized by core splicing components, these interactions alone are not sufficient for splice site selection. There are many *cis*-acting elements, which work together with the core splicing components to promote or discourage the use of splice sites, both constitutive and alternative. Additionally, splice site selection can be affected by various circumstances, including exon and intron length, previous splicing events, pre-mRNA secondary structure, chromatin structure and modifications, and the rate of transcription of the pre-mRNA. All of these factors combine to determine which splice sites will be used and with what frequency.

1.1.2.1 Regulatory Proteins and Sequences

Traditional predictions of splice site usage based solely on the sequence strength of the key splicing elements have not proven to be very accurate (Roca *et al.*, 2013). This is partially because these predictions ignored the wealth of regulatory sequences throughout the pre-mRNA, which can have profound effects on splice site selection. These sequences can be grouped into four main categories based upon their location and behaviour. Sequences that promote splice site usage by recruiting activator proteins are known as enhancers. They are found most often in exons (exonic splicing enhancers – ESEs), but in some cases are also found in introns (intronic splicing enhancers – ISEs). Conversely, sequences that recruit repressor proteins and inhibit

INTRODUCTION

splice site usage are typically found in introns (intronic splicing silencers – ISSs) and, less frequently, in exons (exonic splicing silencers – ESSs) (Zhang & Chasin, 2004). ESEs have been found to be enriched in exons with weak splice sites (Fairbrother *et al.*, 2002). Interestingly, intronic sequences are much more conserved around alternatively used exons when compared to constitutive exons, suggesting that intronic regulatory elements are important for alternative splicing (Sorek & Ast, 2003). Unfortunately, large-scale identification and prediction of intronic regulatory elements is still rather lacking in comparison to that of exonic regulatory elements.

Much research has gone into identifying the activator and repressor proteins that bind these *cis*-regulatory elements and into determining the mechanisms by which they influence splice site selection. The most prevalent and well-studied group of activator proteins are the serine/arginine-rich proteins (SR proteins), and they often work antagonistically with the most common group of repressor proteins, heterogeneous nuclear ribonucleoproteins (hnRNPs) (Eperon *et al.*, 2000; Zhu *et al.*, 2001; Expert-Bezançon *et al.*, 2004; Okunola & Kraine, 2009). However, there are also other important regulatory proteins in addition to SR proteins and hnRNPs, such as TIA-1, Sam68, and the FOX and NOVA proteins (Chen & Manley, 2009).

SR proteins are characterized by C-terminal serine/arginine-rich domains (RS domains) and one to two N-terminal RNA recognition motifs (RRMs) (Long & Cáceres, 2009). In addition to splicing, they have been found to be involved with a wide range of cellular processes related to gene expression, including RNA transport, nonsense-mediated decay (NMD), translation and potentially transcriptional elongation (Zhong *et al.*,

INTRODUCTION

2009). SR proteins have been shown to have several roles in splice site selection, especially for exons with weak splice sites. Stabilization of U2AF35 binding to the 3'ss via interaction of the RS domains of both U2AF35 and the SR protein is one important mechanism, especially when the polypyrimidine tract is weak (Lavigneur *et al.*, 1993; Tian & Maniatis, 1993; Wu & Maniatis, 1993; Staknis & Reed, 1994; Wang *et al.*, 1995). Additionally, SR proteins have been shown to promote U1 snRNP binding, thus influencing 5'ss selection (Eperon *et al.*, 1993; Tarn & Steitz, 1994; Zahler & Roth, 1995). Studies on the well-known SR protein SRSF1 indicate that this interaction is likely to be mediated by the RRM domains of the U1 snRNP associated protein U1 70K and that of the SR protein (Eperon *et al.*, 2000; Cho *et al.*, 2011; Roca *et al.*, 2013). The ability of the RRM domain to participate in this interaction is regulated by phosphorylation of its RS domain, where hypophosphorylation leads to inhibitory binding of the RRM domain by the RS domain (Cho *et al.*, 2011). SR proteins have also been shown to play a role in later spliceosome assembly via direct contacts between the RS domain and RNA duplexes formed between either the U2 snRNA and the BPS or the U6 snRNA and the 5'ss (Shen & Green, 2006; Shen *et al.*, 2004). SR proteins are important not only for alternative splicing but also for constitutive splicing events.

HnRNPs are a relatively diverse group of RNA-binding proteins involved in nucleic acid metabolism. In addition to their role in the regulation of splicing, they have also been shown to be involved in transcription, mRNA trafficking and the regulation of translation (Han *et al.*, 2010). HnRNPs contain RRM domains and a variety of auxiliary domains, such as the glycine-rich RGG domain. Their role in the regulation of splicing is typically an inhibitory one, which is achieved in two main ways. The first involves

INTRODUCTION

stabilization of U1 snRNP binding to such an extent that it halts the progression of spliceosome assembly at A complex. This mode of action has been observed for polypyrimidine tract binding protein (PTB), also known as hnRNP I (Sharma *et al.*, 2011). The second mechanism of inhibition involves the use of steric hindrance, either by physically blocking the binding of other splicing factors (Tange *et al.*, 2001; Saulière *et al.*, 2006; Spellman & Smith, 2006; De Araújo *et al.*, 2009) or by looping out sections of RNA (Damgaard *et al.*, 2002; Nasim *et al.*, 2002; Sharma *et al.*, 2005). Cooperative binding of multiple units of the same hnRNP protein has been observed in some of these cases and helps to facilitate this type of repression (Eperon *et al.*, 2000; Zhu *et al.*, 2001; Spellman & Smith, 2006; Kashima *et al.*, 2007). Of course there are exceptions to any rule, and there have been a few cases where hnRNPs have had a positive effect on splicing. For example, hnRNP H is known to stimulate exon inclusion when bound to ISEs (Chou *et al.*, 1999; Han *et al.*, 2005). The absolute and relative concentrations of regulatory factors such as hnRNPs and SR proteins are very important in splice site selection and, therefore, the control of alternative splicing (Mayeda *et al.*, 1993; Zahler *et al.*, 2004).

1.1.2.2 5' Splice Site Recognition

In the early stages of spliceosome assembly, the 5'ss is recognized through base pairing to U1 snRNP. Initially, this interaction was thought to be the main mode of selection, indicating that superior complementarity to U1 snRNP would lead to use of a particular splice site (Lerner *et al.*, 1980; Rogers & Wall, 1980). However, it soon became clear that many other sequences and splicing factors were involved in 5'ss selection, allowing for the use of 'weaker' 5' splice sites in addition to or in lieu of

INTRODUCTION

'stronger' ones (Treisman *et al.*, 1983; Wieringa *et al.*, 1983). In fact, it has recently been found that there are over 9,000 5'ss variants in the human genome, and their recognition can be aided by the bulging out of nucleotides to create greater complementary to the U1 snRNP (Roca *et al.*, 2012). In addition to the base pairing of U1 snRNA to the 5'ss, interaction of the U1 snRNP-associated U1C protein with the highly conserved GU nucleotides of the 5'ss is also important for the recognition of the 5'ss by the U1 snRNP (Du & Rosbash, 2002; Pomeranz Krummel *et al.*, 2009).

Although U1 snRNP binding is clearly not the only factor involved in 5'ss selection, it is still typically very important. There have been a few cases, however, where U1 snRNA binding to the 5'ss has not been necessary for recognition (Fukumura *et al.*, 2009; Raponi *et al.*, 2009). For example, in the absence of functional U1 snRNP, splice site usage was recovered by increased concentrations of SR proteins (Crispino *et al.*, 1994; Tarn & Steitz, 1994). In this situation, 5'ss selection has been shown to correlate with complementarity to the U6 snRNP, which replaces U1 binding at the 5'ss in the later stages of spliceosome assembly and forms part of the catalytic core of the spliceosome (Crispino & Sharp, 1995). However, these cases make up the exception, not the rule.

Other factors such as pre-mRNA secondary structure and regulatory proteins can influence 5'ss usage. Recognition of 5'ss sequences by the U1 snRNP can be hindered by secondary structures in pre-mRNAs, and there have been many examples of secondary structures influencing alternative splicing events (Eperon *et al.*, 1986; Eperon *et al.*, 1988; Solnick, 1985; Jin *et al.*, 2011). Regulatory proteins, such as SR proteins and hnRNPs, often affect 5'ss selection by either promoting or disrupting U1

INTRODUCTION

snRNP binding (Roca *et al.*, 2013). ESEs and tethered RS domains mimicking SR protein-bound ESEs, when located between alternative 5' splice sites, have been found to promote the usage of the intron-proximal 5'ss (Bourgeois *et al.*, 1999; Gabut *et al.*, 2005; Spena *et al.*, 2006; Wang *et al.*, 2006; Erkelenz *et al.*, 2013). Interestingly, introduction of an ESE at the 5' end of a construct containing two identical alternative 5' splice sites resulted in promotion of the nearest site, which was the intron-distal site (Lewis *et al.*, 2012). Alternatively, the repressor proteins hnRNP A1 has been shown to influence 5'ss selection in favor of the intron-distal site by disruption of U1 snRNP binding (Mayeda & Krainer, 1992; Mayeda *et al.*, 1993; Yang *et al.*, 1994; Eperon *et al.*, 2000). Positioning of the binding sites for these regulatory proteins is clearly important and may explain why some activator proteins have been shown to act as repressors and some repressors as activators (Erkelenz *et al.*, 2013).

Interestingly, U1 snRNP binding at a particular 5'ss does not ensure usage of that site by the spliceosome. This is clearly illustrated by a recent study in which single molecule methods were employed to show that a single pre-mRNA containing two strong 5' splice sites was bound by two U1 snRNPs under E complex conditions (ATP depletion) and that only one remained under A complex conditions (Hodson *et al.*, 2012). There are several potential factors that could affect how splice site recognition by U1 snRNP ultimately results in the use of a 5' splice site, including binding affinity, positional effects and concentrations of U1 snRNP, as well as regulatory factors. Further research is necessary to determine which mechanism(s) are responsible for the final commitment to a particular 5'ss choice (Roca *et al.*, 2013).

INTRODUCTION

1.1.2.3 3' Splice Site Recognition

Recognition of the 3'ss involves three key pre-mRNA sequences: the polypyrimidine tract (PPT) near the 3'ss, the branch point sequence (BPS) located just upstream of the the polypyrimidine tract (PPT) and the well-conserved AG dinucleotide at the 3'ss (Reed, 1989). The BPS is recognized by the SF1 protein, which is later displaced by the U2 snRNP (Kramer & Utans, 1991; Berglund *et al.*, 1998; Wahl *et al.*, 2009). The PPT and AG dinucleotide are recognized by the U2 snRNP auxiliary factor (U2AF), consisting of a larger subunit U2AF65, which binds the PPT, and a smaller subunit U2AF35, which recognizes the AG dinucleotide (Zamore & Green, 1989; Wu *et al.*, 1999). These three sequences are the key components for 3'ss recognition. As seen with the 5'ss (Section 1.1.2.2), sequence strengths and regulatory factors play an important role in 3'ss selection.

Binding of the PPT by U2AF65 is typically the most important factor in 3'ss recognition. U2AF65 binds preferentially to stretches of uridines with its first two RRM domains (Singh *et al.*, 1995; Banerjee *et al.*, 2003; Banerjee *et al.*, 2004), stabilizes SF1 binding to the BPS with its third RRM (Selenko *et al.*, 2003) and promotes U2 snRNP base-pairing to the BPS via its RS domain (Valcárcel *et al.*, 1996). However, U2AF65 binding and U2 snRNP binding are not always directly linked (De Araújo *et al.*, 2009). The BPS and AG dinucleotide as well as enhancer sequences become more important when the PPT is short and/or low in uridine content, causing weak binding of U2AF65 (Reed, 1989; Roecigno *et al.*, 1993; Coolidge *et al.*, 1997). For example, U2AF65 binding to weak PPT sequences has been found to be dependent on U2AF35 binding and enhancer sequences (Tian & Maniatis, 1993; Wang *et al.*, 1995; Zuo & Maniatis, 1996;

INTRODUCTION

Henscheid *et al.*, 2008). Additionally, shortening of a strong polypyrimidine tract can make recognition by U2AF65 enhancer-dependent (Tian & Maniatis, 1992; Graveley & Maniatis, 1998; Zhu & Krainer, 2000). Likewise, enhancement of U2AF65 binding was found to allow for 3'ss recognition of previously enhancer-dependent sites in the absence of the SR proteins that would typically be required for recognition (Tian & Maniatis, 1994; Lorson & Androphy, 2000; Graveley *et al.*, 2001). SR proteins are generally considered to promote 3'ss recognition by stabilizing U2AF35 binding via the RS domains of both the SR proteins and U2AF35 (Lavigneur *et al.*, 1993; Tian & Maniatis, 1993; Wu & Maniatis, 1993; Wang *et al.*, 1995). U2AF65 binding would then be promoted via its interaction with U2AF35 (Zamore & Green, 1989; Kielkopf *et al.*, 2001; Henscheid *et al.*, 2008). Another potential role of SR proteins in 3'ss recognition is to block the binding of various repressor proteins, such as hnRNP A1, Sam68 and PTB, which have been shown to antagonize U2AF65 binding (Kan & Green, 1999; Eperon *et al.*, 2000; Saulière *et al.*, 2006; Paronetto *et al.*, 2011; Tavanez *et al.*, 2012). Also, SR proteins have been shown to stabilize U2 snRNA base-pairing to the BPS via RS domain contacts, similar to the way in which U2AF65 is thought to stabilize U2 snRNP binding (Shen *et al.*, 2004; Shen & Green, 2004). Interestingly, it has been found that a strong BPS can be required to compensate for a weak PPT, while the BPS is not required for recognition of a strong PPT (Reed, 1989; Smith *et al.*, 1989; Coolidge *et al.*, 1997). This flexibility allows for a wide variety of PPT and BPS sequences to be recognized and provides further ways in which regulatory proteins can alter splice site choices (Mount, 1982; Shapiro & Senapathy, 1987).

INTRODUCTION

1.1.2.4 Exon Definition vs. Intron Definition

The ability of the spliceosome to so precisely define exons and introns is rather amazing, considering the length and variation of the sequences involved. The length of exons and introns can vary significantly between different organisms and even within a single pre-mRNA. Humans and other higher eukaryotes tend to have longer introns and shorter exons, while the opposite is observed for lower eukaryotes (Zhang, 1998; Hawkins, 1988; Sakharkar *et al.*, 2005). The relative length of introns and exons has been shown to affect the way in which exons and introns are defined. When exons are short and introns are long, the exons are mapped out first by the basic factors recognizing the 5' and 3' splice sites (U1 snRNP, U2AF) and associated factors, such as SR proteins (Figure 1.5) (Robberson *et al.*, 1990; Berget, 1995; Fox-Walsh *et al.*, 2005). After 'exon definition' occurs, the splicing factors must still make the transition to communication across the introns; however, this process is poorly understood at present (De Conti *et al.*, 2013). When introns are short, then communication of the splicing factors can more easily occur across the intron in the first instance, a process known as intron definition (Figure 1.5) (Lang & Spritz, 1983; Berget, 1995; Sterner *et al.*, 1996). Ultimately, the mechanism of splicing is the same in both cases. In fact, exon and intron definition could both occur for different splicing events within the same pre-mRNA.

Figure removed for copyright reasons.
Please see hard copy version of this thesis
or reference(s) below for the image.

Figure 1.5: Diagram of Exon and Intron Definition (De Conti *et al.*, 2013)

Exon definition (top) is favored when introns are long, while intron definition (bottom) is favored when introns are small.

1.1.2.5 Types of Alternative Splicing

Alternative splicing is an important process for proteome diversity, whereby individual pre-mRNA transcripts are spliced in different patterns to create multiple unique mRNAs. There are several different types of alternative splicing that can be employed to create this diversity. Exon skipping is the most obvious type of alternative splicing, which involves the complete omission of an exon and its flanking introns (Figure 1.6A). This type of alternative splicing is common in higher eukaryotes, but rare in lower

INTRODUCTION

eukaryotes (Sugnet *et al.*, 2004; Alekseyenko *et al.*, 2007; Kim *et al.*, 2007). In some pre-mRNAs, sets of exons are used mutually exclusively (Figure 1.6E). This strategy is quite useful for creating proteins with one particular type of variable domain. For example, the Dscam pre-mRNA contains three cassettes of variable exons, which are used mutually exclusively to produce thousands of different protein combinations, with each final protein containing three variable Ig domains (Wojtowicz *et al.*, 2007). Some exons contain more than one potential 3' or 5' splice site. Alternative use of these splice sites gives rise to two of the other main types of alternative splicing, alternative 3'ss selection (Figure 1.6B) and 5'ss selection (Figure 1.6C). In some alternative splicing cases, introns are retained in the final mRNA (Figure 1.6D). This type of alternative splicing is rare in humans and higher eukaryotes, but is quite common in lower eukaryotes where intron lengths tend to be shorter (Sugnet *et al.*, 2004; Alekseyenko *et al.*, 2007; Kim *et al.*, 2008). Other rare events include variable promoters or polyadenylation sites at the extreme 5' or 3' ends of the pre-mRNA (Figure 1.6F and G) and *trans*-splicing, which involves the joint splicing of two separate RNA molecules (Suzuki *et al.*, 2001; Beadoing *et al.*, 2000; Murphy *et al.*, 1986; Flouriot *et al.*, 2002). All of these types of alternative splicing illustrate the flexibility of the spliceosome and highlight its impressive role in creating diversity.

Figure removed for copyright reasons.
Please see hard copy version of this thesis
or reference(s) below for the image.

Figure 1.6: Diagram of the Different Types of Alternative Splicing (Keren *et al.*, 2010)

(A) Exon skipping – An entire exon with flanking intronic sequences is omitted by the spliceosome. **(B)** Alternative 3' splice site selection – Multiple 3' splice sites within an exon can be used to generate differently spliced products. **(C)** Alternative 5' splice site selection – Multiple 5' splice sites within an exon can be chosen, resulting in alternatively spliced products. **(D)** Intron retention – This form of alternative splicing is common in lower eukaryotes, where intron lengths tend to be shorter. **(E)** Mutually exclusive exons – A set of two or more exons that are included mutually exclusively into the final mRNA. This type of alternative splicing can be used, for example, to generate protein isoforms containing one particular variable domain. **(F)** Alternative promoters – Multiple promoter sites allows for variation at the 5' end of the mRNA. **(G)** Alternative polyadenylation – Alternative polyadenylation can be achieved as depicted above with variable 3' exons or by various other mechanism. The study of polyadenylation is a field of research in itself.

1.2 Alternative Splicing and Disease

Pre-mRNA splicing is a very complex process involving many levels of regulation, ultimately leading to the location and union of exons across vast expanses of intronic sequences. Equally impressively, the catalytic macromolecular machine responsible for this amazing process, the spliceosome, must be reassembled on every pre-mRNA transcript, unlike other cellular machinery, such as ribosomes, which process many transcripts before being recycled (Wang & Burge, 2008; Wahl *et al.*, 2009). Given this complexity and the fact that an estimated 94% of human genes are alternatively spliced, it is no wonder that up to 50% of disease-causing mutations are thought to affect splicing (Wang *et al.*, 2008; Pan *et al.*, 2008; López-Bigas *et al.*, 2005).

Research to date has discovered many diseases, both genetic and acquired, involving problems related to splicing. As splicing is such a complex process, there are many different ways that mutations could result in adverse effects. Mutations in either the 5' or 3' splice sites are possibly most obvious and make up roughly 10% of disease-causing mutations (Krawczak *et al.*, 2007). That is approximately one fifth of the estimate for all splice-altering mutations resulting in disease (López-Bigas *et al.*, 2005; Krawczak *et al.*, 2007). Mutations affecting recognition or function of the branch point sequence, the other fundamental splicing sequence vital to the first of the two transesterification reactions responsible for intron excision, could also be devastating. As alternative splicing, and splicing in general, involves the use of many different regulatory sequences in exons (ESEs, ESSs) and introns (ISEs, ISSs), the whole pre-mRNA could be seen as a mutation minefield, where one point mutation could

INTRODUCTION

eliminate or create enhancing and silencing sequences, thus changing the natural patterns of exon skipping or inclusion.

The function of the disrupted protein will obviously affect the nature of the resultant disease. Alterations in behaviour or expression levels of splicing regulatory proteins or machinery could have widespread effects. Many cancers have been shown to involve changes to the alternative splicing or expression of tumor suppressors and proto-oncogenes (Skotheim & Nees, 2007; Ghigna *et al.*, 2008; Kaida *et al.*, 2012). For example, the splicing activator protein SRSF1 is actually considered to be a proto-oncogene due to the fact that its overexpression in many types of cancer leads to all sorts of changes in alternative splicing, ultimately resulting in or promoting a cancerous cell state (Karni *et al.*, 2007). Just one of the known effects of SRSF1 overexpression is increased cell motility and metastasis due to SRSF1 induced skipping of exon 11 in the proto-oncogene *Ron* (Ghigna *et al.*, 2005). Another example, which illustrates the power of a single change to the alternative splicing of a gene, involves the apoptotic regulator gene *Bcl-X*. It has two alternative splice products: *Bcl-X_S*, which is pro-apoptotic, and *Bcl-X_L*, which is anti-apoptotic and overexpressed in many cancers (Ma *et al.*, 2010). The balance between these two isoforms is controlled by many splicing regulatory proteins, such as hnRNP K and Sam68, further illustrating how important and delicate the balance of splicing factors is to the maintenance of healthy cells (Paronetto *et al.*, 2007; Revil *et al.*, 2009).

Even though splicing is a very intricate process, not all mutations will be as severe as others. Some may only disrupt the balance slightly, resulting in production of a lower

INTRODUCTION

percentage of the correct splice product rather than complete loss. Also, even if an exon is skipped out in 100% of transcripts or a premature stop codon is generated, there is still a chance that the protein produced will retain some functionality. As different isoforms are produced in different abundances dependent upon cell type and developmental stage, it is important to understand the natural patterns of splicing prior to investigation of the effects of mutations. Once these things are understood, it is possible to develop specific therapeutic strategies.

1.2.1 Therapeutic Strategies to Influence Alternative Splicing

As the number of understood links between splicing and disease grows, so too does the potential for the development of splice-altering therapeutic strategies. As there are many different ways that disease-causing mutations can affect the splicing or subsequent translation of pre-mRNAs, many different strategies of intervention could be imagined and explored. This section covers the two main strategies for manipulation of alternative splicing; however, there are other potentially useful RNA based therapeutic strategies targeting other aspects of mRNA processing (Cooper *et al.*, 2009). Practical considerations such as cost, ease of delivery, and specificity have a great impact on the viability of any promising potential therapy. This is certainly the case for splice-altering therapies involving small molecules or antisense oligonucleotides, both having their own set of advantages and disadvantages.

1.2.1.1 Small Molecules

Splicing regulatory proteins, such as SR proteins and hnRNPs, are prime targets for small molecules as they influence most splicing events. The balance between these

INTRODUCTION

proteins often has the power to determine the outcome of a splicing event, so small molecules that influence their expression levels or behaviour could have significant effects on alternative splicing. Small molecule inhibitors of SR proteins and the kinases that regulate their phosphorylation, and therefore function, have already been identified and shown promise (Muraki *et al.*, 2004; Soret *et al.*, 2005; Fukuhara *et al.*, 2006; Graveley, 2005). The obvious downside to this approach is the lack of selectivity and potential for many off-target effects. However, in some cases a widespread reduction in activity may be ideal. For example, in cancers where the SR protein SRSF1 is overexpressed, affecting the alternative splicing of many different pre-mRNAs, inhibition using small molecules could be appropriate and help to return the alternative splicing patterns to a more normal state (Wang & Manley, 1995; Karni *et al.*, 2007). For example, the pro-metastatic alternative splicing event in Ron, which is promoted by SRSF1, has been shown to be corrected by the use of indole derivatives, which inhibit SRSF1 activity (Ghigna *et al.*, 2010). Small molecules targeting other splicing-related proteins, such as SF3b, which is part of the U2 snRNP, have shown promise as anti-tumor agents, even progressing to clinical trials (Fan *et al.*, 2011; Kaida *et al.*, 2007; Kong & Yamori, 2012; Kotake *et al.*, 2007). Other benefits of small molecule strategies, compared to antisense oligonucleotide strategies for example, include easier delivery and lower costs.

1.2.1.2 Antisense Oligonucleotides

Antisense oligonucleotide based strategies for splice-altering therapies are becoming increasingly more promising and represent a more targeted, specific approach in comparison to small molecule strategies (Eperon, 2012). The use of antisense

INTRODUCTION

oligonucleotides to alter splicing patterns is different from the traditional gene-silencing application of antisense oligonucleotides in that the goal is not to induce RNase H digestion. The mRNA needs to remain intact, so chemically modified oligonucleotides are used, which are resistant to RNase H digestion. Figure 1.7 shows some of the main chemical modifications, which can provide various benefits in addition to RNase H resistance, such as increased binding affinity, resistance to nuclease degradation and improved uptake (Deleavey & Damha, 2012). Chemical modifications can have significant effects on the functionality of splice-altering oligonucleotides, so it is important to take this into account during development (Owen *et al.*, 2011). In fact, recently a group discovered that an oligonucleotide sequence, which stimulated a splicing event by blocking a repressor protein, had an inhibitory effect after 2' F modification. This was found to be due to recruitment of the inhibitory interleukin enhancer-binding factor 2 and 3 (ILF2/3) complex by the 2' F oligonucleotide, which was not recruited by the same sequence when 2' MOE modified (Rigo *et al.*, 2012). This finding, whilst potentially providing a new strategy for the induction of exon skipping, also stresses the importance of oligonucleotide chemistry.

INTRODUCTION

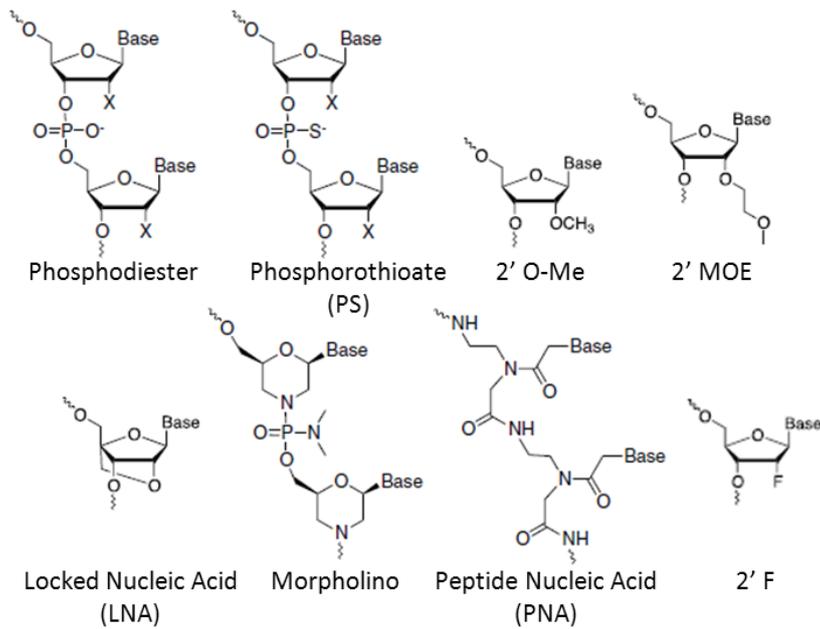


Figure 1.7: Types of chemical modifications for oligonucleotides

Phosphodiester bonds are used link together bases in natural DNA and RNA. Various chemical modifications have been developed to modify the properties of synthetic oligonucleotides. Because RNA oligonucleotides are prone to degradation and RNase H digestion, modifications such as 2' OMe and 2' MOE, provide stability. Phosphorothioate modifications are also used to provide protection and stability. Other modifications such as morpholino and peptide nucleic acid affect the entire backbone structure and can provide other benefits such as improved delivery *in vivo*. The locked nucleic acid modification affects the secondary structure of the oligonucleotide, locking it into a rigid position that favors annealing.

Figure 1.8 illustrates the four main ways that antisense oligonucleotides are used to alter splicing patterns, all of which involve manipulation of the binding of regulatory proteins. In order to induce exon inclusion, antisense oligonucleotides can either be used to block the binding of a repressor protein (Figure 1.8A) or recruit an activator protein (Figure 1.8B). Stimulation of splicing using targeted bifunctional oligonucleotide enhancers of splicing (TOES), which have an annealing region for targeting to the correct location and a non-annealed tail containing an enhancer sequence for recruitment of an activator protein, is a strategy that was first developed in 2003 (Skordis *et al.*, 2003). Bifunctional oligonucleotides can also be used to induce

INTRODUCTION

exon exclusion if the tail is composed of a silencer sequence that recruits a repressor protein (Figure 1.8D). Exon skipping can also be achieved by blocking of an enhancer sequence with an antisense oligonucleotide (Figure 1.8C). The ability to induce exon exclusion has the potential to be very powerful, especially for diseases where mutations have introduced premature stop codons or shifted the reading frame, as removal of one damaged exon could allow for the production of a protein that still retained some functionality. An antisense oligonucleotide that causes skipping of an exon containing a premature stop codon is in phase 2 clinical trials for the treatment of Duchene muscular dystrophy (Kinali *et al.*, 2009; Goemans *et al.*, 2011; Cirak *et al.*, 2011). Also in clinical trials is an antisense oligonucleotide that blocks a silencer sequence in the SMN2 gene, providing a promising therapy for spinal muscular atrophy (ClinicalTrials.gov Identifier NCT01494701).

INTRODUCTION

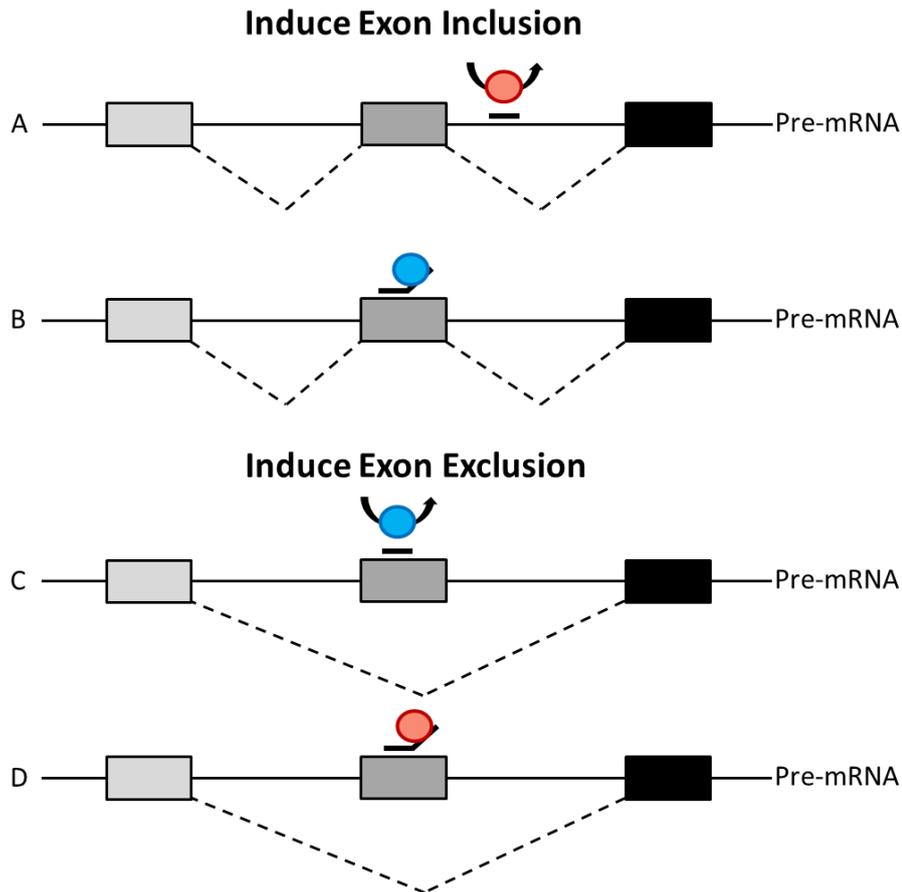


Figure 1.8: Splice-Altering Antisense Oligonucleotide Strategies

The main antisense oligonucleotide-base strategies for the manipulation of alternative splicing involve either the blocking or recruitment of regulatory proteins using RNase H-resistant ASOs. There are two ways to induce exon inclusion: **(A)** anneal over a silencing sequence, thus blocking the binding of a repressor protein or **(B)** recruit an activator protein using a bifunctional oligonucleotide with an enhancer tail. Likewise, there are two ways to induce exon skipping: **(C)** block an enhancer sequence, thus preventing the binding of an activator protein or **(D)** recruit a repressor protein using a bifunctional oligonucleotide with a silencer tail.

As all of the sequences and proteins that regulate alternative splicing are not fully understood, finding the best annealing location for therapeutic antisense oligonucleotides requires some effort or good luck. This, as well as the cost of synthesis of these oligonucleotides, makes them more expensive as potential therapeutics than small molecules. However, they have the potential to be much more specific and reduce or eliminate the many potential side effects associated with small

INTRODUCTION

molecule treatments. Efficient delivery into the right tissues and cellular compartments is also a challenge facing some antisense strategies, especially those requiring delivery across the blood-brain barrier. The use of viral vectors capable of transporting materials across the blood-brain barrier is one way that this could be overcome, but this strategy requires further optimization (Baughan *et al.*, 2006; Geib & Hertel, 2009; Donnelly & Boulis, 2012). Chemical techniques, such as oligonucleotide modifications (Zhou *et al.*, 2013), specialized PEG/PEI capsules for traversing the blood-brain barrier (Vinogradov *et al.*, 2004), triggered-release cages (Venkatesh *et al.*, 2009), and the use of various types of nanoparticles (Yang *et al.*, 2009; Rimessi *et al.*, 2009; Kim *et al.*, 2011), are also potential solutions to the issue of delivery. Despite these challenges, antisense oligonucleotide base splice-altering therapies are very promising. With splicing being at the root of so many genetic and acquired diseases, the understanding, optimization and use of these strategies can only improve and increase over time.

1.3 Spinal Muscular Atrophy

Spinal muscular atrophy (SMA) is the leading genetic cause of infant death, with one in every 6,000 babies born being affected (Lunn & Wang, 2008; Pearn, 1980; Feldkötter *et al.*, 2002). It is characterized by degeneration of motor neurons, caused by greatly reduced levels of the ubiquitously expressed SMN protein, especially in the spinal cord (Lunn & Wang, 2008; Covert *et al.*, 1997). SMN is important for the biogenesis of snRNPs, has been found to be associated with splicing regulatory proteins, and plays a role in the formation of spliceosomal complex E (Mourelatos *et al.*, 2001; Makarov *et al.*, 2012). SMN is involved in a range of other cellular activities beyond those related

INTRODUCTION

to splicing, including transcription, apoptosis, RNA stability and axonal RNA trafficking (Burghes & Beattie, 2009). The loss of two key roles of the SMN protein: snRNP biogenesis and motor axon growth/function, have been independently implicated in the progression of spinal muscular atrophy (Carrel *et al.*, 2006; Coady & Lorson, 2011). To fulfill its role in snRNP biogenesis, SMN functions in the SMN complex, comprised of 9 factors: SMN, Gemins 2–8 and STRAP (aka UNR-IP) (Otter *et al.*, 2007). Restoration of SMN protein levels is vital to the treatment of SMA.

The reduction in SMN protein associated with SMA is typically caused by a homozygous deletion of the SMN1 gene, which produces SMN; however, in a few cases mutations are to blame (Wirth, 2000; Hamilton & Gillingwater, 2013; Monani, 2005). Fortunately, all SMN patients have at least one copy of a very similar gene, SMN2, which is capable of making functional SMN protein; however, silent mutations cause mis-splicing of 90% of pre-mRNA transcripts, leading to the production of a non-functional truncated protein that is rapidly degraded (Le *et al.*, 2000; Monani *et al.*, 1999; Cartegni *et al.*, 2006; Lorson *et al.*, 1998). The 10% of functional SMN produced is still beneficial, leading to a reduction in severity of disease that is correlated with increasing copy numbers of the SMN2 gene (Feldkötter *et al.*, 2002; McAndrew *et al.*, 1997; Taylor *et al.*, 1998). As SMN2 pre-mRNA has the potential to code for functional SMN protein when spliced correctly, it is an ideal target for splice-altering therapies. Knowledge of the splicing regulatory elements of SMN1 and SMN2 is necessary for the development of targeted therapeutic strategies.

1.3.1 Splicing of SMN1 and SMN2

The SMN1 and SMN2 genes are greater than 99% homologous, having only a few nucleotides different between them (Skordis *et al.*, 2001). However, a change from C to T in SMN2 exon 7 is critical to the splicing of SMN2, leading to skipping of exon 7 in SMN2 (Lorson *et al.*, 1999; Monani *et al.*, 1999). Two mechanisms have been proposed to explain the effect that this mutation at the 6th nucleotide of exon 7 has on the splicing of SMN2. One model suggests that disruption of an exonic splicing enhancer (ESE), which binds the activator protein SRSF1 is responsible (Cartegni & Krainer, 2002; Cartegni *et al.*, 2006). Another model suggests that the mutation produces an exonic splicing silencer (ESS), which is bound by hnRNP A1 and/or Sam68 (Kashima & Manley, 2003; Kashima *et al.*, 2007; Pedrotti *et al.*, 2010). In reality, these models are likely to both be valid (Cartegni *et al.*, 2006). In fact, SMN2 has been shown to have a reduced level of U2AF65 binding to the 3' splice site of intron 6, compared to SMN1, and a reduced capacity for recruitment of the U2 snRNP, which is consistent with both models as SRSF1 promotes, while hnRNP A1 antagonizes, binding of these factors (De Araújo *et al.*, 2009). Although in this case, another enhancer region in exon 7, known to bind Tra2- β , was also found to be necessary for U2 recruitment. Given the dramatic impact of this C6T mutation site on SMN2 splicing, it would clearly make a good therapeutic target.

There are also splicing elements common to both SMN1 and SMN2 that could be exploited to alter SMN2 splicing. An ESE in the middle of exon 7, which is bound by Tra2- β , has been shown induce SMN2 exon 7 inclusion under conditions of over-expression of Tra2- β (Hofmann *et al.*, 2000). Alternatively, an intronic splicing silencer

INTRODUCTION

found near the 5' end of intron 7 (ISS-N1) has been shown to have a strong negative effect on exon 7 inclusion, and blocking of this site with an antisense oligonucleotide resulted in increased SMN2 exon 7 inclusion (Singh *et al.*, 2006). Another ISS found in intron 6 has also been found to inhibit exon 7 inclusion (Miyajima *et al.*, 2002). A weak 5' splice site containing an inhibitory secondary structure has also been found to influence the splicing of exon 7 (Singh *et al.*, 2004; Singh *et al.*, 2007). Understanding the key splicing features and regulatory elements of both SMN1 and SMN2 is essential for the development of targeted therapeutic strategies for SMA.

1.3.2 Therapeutic Approaches

The fact that spinal muscular atrophy affects so many babies provides a clear incentive for the development of therapies. This, combined with the understanding of the cause of the disease, loss of SMN1, and identification of a potential substitute, the SMN2 gene, makes SMA such an appealing therapeutic target. Mouse models of SMA with varying degrees of severity have been created for the study of potential therapies prior to clinical trials in humans (Bebbee *et al.*, 2012). Many different potential therapeutic strategies have been developed, each with their own strengths and weaknesses.

Various small molecule strategies have shown some promise, including HDAC inhibitors (Kernochan *et al.*, 2005; Avila *et al.*, 2007; Garbes *et al.*, 2009; Hauke *et al.*, 2009), a readthrough-inducing aminoglycoside (Mattis *et al.*, 2006; Mattis *et al.*, 2009), a pH altering Na⁺/H⁺ exchanger inhibitor (Yuo *et al.*, 2008), and various polyphenol botanical compounds that were found to increase levels of SMN protein in SMA patient fibroblasts (Sakla & Lorson, 2008). However, all of these strategies are non-

INTRODUCTION

specific and therefore could have many undesirable off-target effects. For example, the HDAC inhibitor valproic acid was found to actually decrease motor neuron growth and function in a mouse model of SMA, clearly a highly counterproductive side-effect (Rak *et al.*, 2009). One clear benefit of small molecule therapies, however, is the ease of delivery. This is especially important for SMA because of the nature of the disease and the need for treatments to be able to cross the blood-brain barrier.

Viral vectors, such as the self-complementary adeno-associated virus 9 (scAAV9) vector, have shown great promise for the delivery of a replacement SMN1 gene across the blood-brain barrier in mouse models of SMA (Foust *et al.*, 2010; Azzouz *et al.*, 2004; Valori *et al.*, 2010; Dominguez *et al.*, 2011; Passini *et al.*, 2010). Also, viral delivery of expression vectors for targeted antisense oligonucleotides, which are designed to promote exon 7 inclusion in SMN2 pre-mRNA, looks encouraging in mice (Baughan *et al.*, 2006; Geib & Hertel, 2009). However, establishing the safety and delivery of these viral vectors and convincing the general public of their safety are the biggest challenges currently facing these promising methods.

Another strategy involves the transplantation of motor neuronal stem cells derived from embryonic stem cells. Intrathecal injection of these stem cells did lead to a modest improvement in a mouse model for SMA, but further optimization would be required in order to make it a competitive strategy (Corti *et al.*, 2010). Also, this method is quite invasive, and similar approaches have shown potential safety hazards, such as the development of brain tumors (Amariglio *et al.*, 2009).

INTRODUCTION

Some of the most promising therapeutic strategies involve specifically targeting the various splicing regulatory elements of SMN2 in order to induce exon 7 inclusion, thereby restoring the production of functional SMN protein. Typically this involves the use of antisense oligonucleotides, but in one case it involved a synthetic RS domain, designed to mimic the activating role of SR proteins, which was specifically targeted to exon 7 by a tethered antisense moiety (Cartegni & Krainer, 2003). There are two main antisense oligonucleotide strategies for the alteration of SMN2 splicing. The most simple strategy involves blocking of the various silencer elements in SMN2 with an annealed antisense oligonucleotide, thereby allowing exon 7 inclusion by removing repression (Hua *et al.*, 2008; Singh *et al.*, 2009; Hua *et al.*, 2007; Hua *et al.*, 2010; Passini *et al.*, 2011; Porensky *et al.*, 2012; Williams *et al.*, 2009). Another strategy involves actively stimulating exon 7 inclusion using bifunctional oligonucleotides that contain an annealing region for targeting to the exon, which in some cases blocks a silencer site, and a non-annealed tail region, which contains an enhancer sequence for recruitment of positive splicing factors (Skordis *et al.*, 2003; Owen *et al.*, 2011; Baughan *et al.*, 2009; Meyer *et al.*, 2009; Baughan *et al.*, 2006; Osman *et al.*, 2012). Both the blocking antisense oligonucleotides and the bifunctional targeted oligonucleotide enhancers of splicing (TOES) have shown great promise both *in vitro* and *in vivo*. In fact, one of the blocking antisense oligonucleotides named ISIS SMN_{Rx}, which targets ISS-N1, has shown success in phase 1 clinical trials in infants with SMA (ClinicalTrials.gov Identifier NCT01494701) and represents the current most promising strategy for the treatment of spinal muscular atrophy (Rigo *et al.*, 2012). In October 2013, ISIS SMN_{Rx} was administered to the first patient in a small phase 2 clinical trial designed to assess the optimal dosing for a larger phase 3 clinical trial expected to

INTRODUCTION

begin in early 2014 (News-medical.net, 2013). As many antisense oligonucleotides cannot cross the blood-brain barrier, delivery is more difficult. Injections bypassing the blood-brain barrier, although invasive, are a viable option for treatment. However, there are various, less invasive strategies being developed to facilitate this type of delivery. In fact, systemic delivery of a morpholino modified antisense oligonucleotide blocking ISS-N1 into neonatal SMA model mice was found to increase SMN2 exon 7 inclusion in the central nervous system (CNS), indicating that the oligo was able to cross the blood-brain barrier (Zhou *et al.*, 2013). Also, ISIS SMN_{Rx} was found to have superior results in a mouse model of SMA with simple subcutaneous injections compared to direct delivery into the CNS. However, in this case splicing in the CNS was not affected and the result was attributed to higher levels of expression of insulin-like growth factor 1 (IGF1) in the liver (Hua *et al.*, 2011; Rigo *et al.*, 2012). Therefore, it is more difficult to predict if a similar result would be seen in humans, given the differences in liver metabolism between mice and humans. Improving delivery of therapeutic antisense oligonucleotides across the blood-brain barrier is important for diseases involving the central nervous system. The specificity of antisense oligonucleotide splice-altering strategies makes them ideal candidates for therapeutic applications in spinal muscular atrophy and other splicing-related diseases.

In addition to being useful therapeutic agents, antisense oligonucleotides capable of modifying splicing patterns can also be used to study the regulatory mechanisms of alternative splicing. The main objective of this thesis is to investigate the precise mechanism of action of the first bifunctional TOES developed for the treatment of SMA named GGA, using a wide variety of biochemical strategies (Skordis *et al.*, 2003).

1.4 G-Quadruplexes

GGA consists of a 15 nucleotide annealing region and a 23 nucleotide enhancer tail.

The tail sequence of GGA (AGGAGGACGGAGGACGGAGGACA) has recently been shown to be able to form a G-quadruplex structure (Smith *et al.*, manuscript submitted).

When this was discovered, the effect of this structure on the function of GGA as an enhancer of splicing was unknown.

The field of G-quadruplex research is rapidly growing as more and more biological processes are shown to involve G-quadruplexes. Perhaps the most widely-known G-quadruplexes are DNA quadruplexes found at the ends of telomeres where they provide a structural barrier to the telomerase enzyme, thus ensuring eventual cell mortality (Fletcher *et al.*, 1998). DNA G-quadruplexes have also been shown to play a role in the regulation of gene expression at the level of transcription, adding to the list of secondary structures involved in the regulation of transcription (Huppert & Balasubramanian, 2007). G-quadruplex prediction software estimates that there are up to 376,000 potential G-quadruplex forming motifs in the human genome. G-quadruplexes are also associated more heavily with different categories of genes, indicating a regulatory role (Huppert & Balasubramanian, 2005; Todd *et al.*, 2005). For example, G-quadruplexes were found to be over-represented in proto-oncogenes and under-represented in tumor suppressor genes (Eddy & Maizels, 2006). The role of G-quadruplexes in telomere maintenance, as well as their association with oncogenes, makes them prime cancer drug targets. This has led to the development of various G-quadruplex stabilizing ligands that, in addition to being potential therapeutics, could also be useful tools for further investigations into G-quadruplexes (Düchler, 2012).

INTRODUCTION

It is now known that RNA also can form G-quadruplexes, both *in vitro* and *in vivo* (Millevoi *et al.*, 2012; Wieland & Hartig, 2007). In fact, RNA G-quadruplexes form more readily, due to the single stranded nature of RNA, and are stronger than DNA G-quadruplexes (Saccà *et al.*, 2005). RNA G-quadruplexes have been suggested to be involved in a wide range of cellular processes including transcription termination and polyadenylation of mRNA transcripts (Huppert *et al.*, 2008; Zarudnaya *et al.*, 2003), translation initiation (Huppert *et al.*, 2008; Beaudoin & Perreault, 2010; Morris *et al.*, 2010), telomere processing (Gros *et al.*, 2008), mRNA targeting (Subramanian *et al.*, 2011) and alternative splicing (Kostadinov *et al.*, 2006; Kikin *et al.*, 2008). It is easy to envisage how the strong G-quadruplex structures could have inhibitory effects by forming a physical barrier to the binding of proteins. However, the spectrum of G-quadruplex involvement seems to be much more complex than this. A growing number of proteins have been found to either bind G-quadruplex structures or the sequences that can form them (Millevoi *et al.*, 2012; Fry, 2007). From the evidence found thus far, it does seem that G-quadruplexes play a largely negative role in the regulation of translation initiation, but not exclusively (Huppert *et al.*, 2008; Beaudoin & Perreault, 2010; Morris *et al.*, 2010). Also, RNA G-quadruplexes in various pre-mRNAs have been found to have both silencing (Gomez *et al.*, 2004) and enhancing effects on alternative splicing (Marcel *et al.*, 2011; Didiot *et al.*, 2008). It is important that research into the function of G-quadruplexes assesses the effect of surrounding sequences that could alter G-quadruplex formation and takes into account the structural state when various proteins are bound. Unfortunately, this has not always been done studies thus far. Further research is needed to better understand, and therefore potentially

INTRODUCTION

manipulate, the wide variety of RNA and DNA G-quadruplex structures, which are highly prevalent and involved in many different regulatory processes.

Many different types of proteins, including helicases, exoribonucleases, chaperone proteins, transcription factors and splicing regulatory proteins, have been found to be associated with sequences that can form G-quadruplexes (Fry, 2007; Millevoi *et al.*, 2012). Some of these proteins promote the formation and/or stabilization of the G-quadruplex structure, while others promote unfolding. The number of proteins known to be involved with G-quadruplexes will no doubt continue to grow. The wide range of G-quadruplex associated proteins reflects the ubiquitous, yet variable, nature of G-quadruplexes.

1.4.1 G-quadruplex Structure

There is a significant amount of variation in G-quadruplex structures and the sequences that form them. However, there are several key features that define G-quadruplexes. As the name suggests, they are formed by planar stacks of four guanine bases called G-tetrads (Figure 1.9A), which form Hoogsteen hydrogen bonds between each other, with a stabilizing monovalent cation coordinated in the middle of or between the G-tetrads (Figure 1.9B) (GELLERT *et al.*, 1962; Bochman *et al.*, 2012). Once formed, this bonding arrangement is much more stable than typical base pairing and requires either high temperatures or the binding of specific proteins to destabilize it (Simonsson, 2001; Pilch *et al.*, 1995; Harrington *et al.*, 1997; Lane *et al.*, 2008).

Potassium ions confer the greatest level of stability, followed by sodium (Jim *et al.*, 1992). Therefore, the ions present in physiological conditions favor G-quadruplex

INTRODUCTION

formation. G-quadruplexes can vary in the number of stacked G-tetrads, with a minimum of two (Huppert, 2010). Also, the loops between the G-runs can vary in length and sequence, leading to a large number of potential G-quadruplex forming sequences (Patel *et al.*, 2007). The stability of G-quadruplexes is affected by both the number of G-tetrads and the length of the adjoining loops, with longer G-runs and shorter loops providing the most stability (Huppert, 2010). G-quadruplex formation is not limited to single stranded (intramolecular) folding either. Intermolecular G-quadruplexes can form between two separate strands, each contributing two G-runs, or even between four individual strands. The orientation of each strand, parallel verses antiparallel, can also vary, further increasing the complexity of G-quadruplex structures (Simonsson, 2001). It has been shown that individual G-quadruplex motifs are also able to stack together to form higher order complexes (Lu *et al.*, 1992; Sen & Gilbert, 1992). The variability of G-quadruplex structures helps to explain their prevalence throughout the genome and transcriptome, and sparks interest into the types of proteins that bind to these sequences and their roles within the cell.

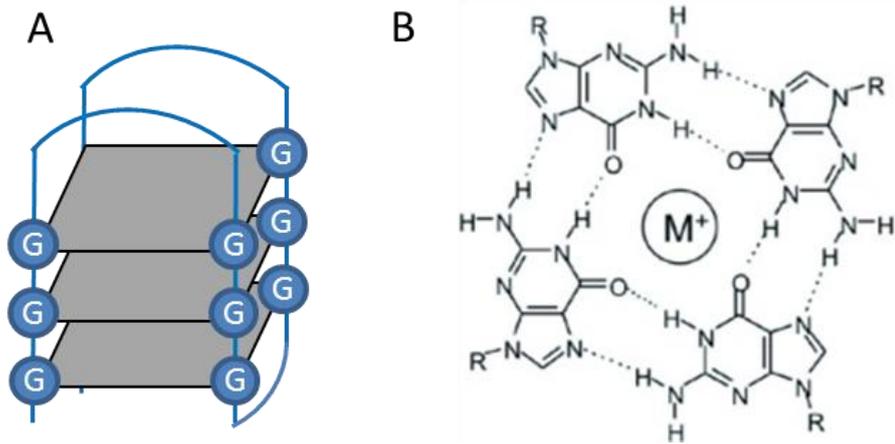


Figure 1.9: Basic G-quadruplex Structure

(A) G-quadruplexes are formed by sequences containing runs of G's with intervening sequences, which become loops. These G-runs assemble into stacks of G-tetrads (grey boxes) with adjoining loops. This drawing of an intramolecular G-quadruplex with three G-tetrads is just an example of a much wider range of possible structural arrangements. (See (Simonsson, 2001) for review) **(B)** A top-down view of a single G-tetrad shows the Hoogsteen hydrogen bonding between the four guanines as well as the coordinated cation in the middle, which provides stability (Huppert & Balasubramanian, 2005).

G-quadruplexes have been shown to have both positive and negative effects on various cellular processes. It was therefore interesting to investigate the possible influence that a G-quadruplex structure could have on the function of the enhancer tail of GGA as part of a wider investigation into the overall mechanism of action of GGA. This knowledge would improve the therapeutic potential of GGA and other TOES.

1.5 Specific Aims

The overall aim of this study was to determine the mechanism of action of a tailed oligonucleotide enhancer of SMN2 exon 7 splicing. This oligonucleotide, named GGA, was first designed and shown to be successful in stimulating exon 7 inclusion in fibroblasts from patients with spinal muscular atrophy in 2003 (Skordis *et al.*, 2003). The same sequence when incorporated into a U7 snRNP and transgenically expressed in SMA model mice provided a remarkable improvement essentially giving them the abilities and longevity of a normal mouse (Meyer *et al.*, 2009). Therefore, the potency and potential of GGA have been clear for some time; however, it was still unclear precisely how this stimulation was achieved. Many different *in vitro* biochemical techniques were employed in this study, including various splicing assays, UV crosslinking, FeBABE experiments, single molecule experiments, electrophoretic mobility shift assays, and biotin affinity purification with photocleavage, an improved technique developed during this study. Although GGA was the main focus of this study, two other antisense oligonucleotide strategies for the stimulation of SMN2 exon 7 splicing were also briefly investigated using *in vitro* splicing. The ultimate goal of understanding the mechanism by which GGA enhances splicing would be to provide insights into how enhancer sequences in general work and how they could best be exploited in future therapeutic oligonucleotide designs.

2 MATERIALS AND METHODS

2.1 *In vitro* Protein Techniques

2.2 *In vitro* RNA Techniques

2.3 Cell Culture and Nuclear Extract Preparation

2.4 Single Molecule Techniques

2.5 Techniques Related to Fe-BABE Experiments

METHODS

2.1 *In vitro* Protein Techniques

2.1.1 Biotin Affinity Purification Methods

2.1.1.1 Biotin Affinity Purification with α -U1 Oligonucleotide

A third of the sample volume of NeutrAvidin® agarose beads (5 μ l, 50% slurry, Thermo Scientific) were pre-washed at 4 °C with 1 ml SDS wash buffer (100 mM Tris-HCl pH 7.5, 1% SDS, 10 mM DTT), followed by 3 x 1 ml buffer (100 mM NaPO₄ pH 7.2, 150 mM NaCl), then resuspended in 1 ml chilled binding buffer (100 mM NaPO₄ pH 7.2, 150 mM NaCl, and either no detergent, 0.01% Triton-X 100 or 0.05% NP-40). Meanwhile, 15 μ l samples containing 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, 250 nM α -U1 2'OMe oligonucleotide (Section 8.2.4) and 40% commercial HeLa nuclear extract (Cilbiotech) were set up and incubated at 30 °C for 20 min. Precipitated proteins were then removed from the samples by centrifugation at 13000 RPM for 1 min before addition of supernatant to the prewashed beads in binding buffer at 4 °C. Binding was allowed to proceed with constant mixing at 4 °C overnight. Then the entire sample was placed onto a column with a frit to retain the beads (Bio-Rad) and washed with 30 ml chilled buffer with salt (100 mM NaPO₄ pH 7.2, 150 mM NaCl, 0.1% Triton-X 100) followed by 5 ml chilled buffer without salt (100 mM NaPO₄ pH 7.2, 0.1% Triton-X 100). Samples were eluted from beads on the column by adding 50 μ l SDS dyes and heating at 90 °C for 5 min. SDS PAGE was used for analysis.

2.1.1.2 Biotin Affinity Purification using a Photocleavable Linker

Samples of 60 μ l, 300 μ l or 1 mL were purified using the following method. A third of the sample volume of NeutrAvidin® agarose resin (50% slurry, Thermo Scientific) was

METHODS

pre-washed at 4 °C with 1 ml SDS wash buffer (100 mM Tris-HCl pH 7.5, 1% SDS, 10 mM DTT), followed by 3 x 1 ml sodium phosphate buffer (100 mM NaPO₄ pH 7.2, 150 mM NaCl, 0.05% NP-40), then resuspended in 1 ml chilled sodium phosphate buffer. Meanwhile, samples containing 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, 250 nM GGA PC Bio (Section 8.2.1) and 40% commercial HeLa nuclear extract (Cilbiotech) were set up and incubated at 30 °C for 20 min in the dark. Samples were protected from light from this point until after UV cleavage step. Precipitated proteins were then removed from the samples by centrifugation at 13000 RPM for 1 min before addition of the supernatant to the prewashed beads in binding buffer at 4 °C. Binding was allowed to proceed with constant mixing at 4 °C for 2 hours or overnight. Then the entire sample was placed onto a column with a frit to retain the beads (Bio-Rad) and washed with 40 ml chilled sodium phosphate buffer followed by 40 ml chilled sodium phosphate buffer without the salt. Cleavage of the photocleavable linker of GGA PC Bio was achieved by application of 365 nm UV light (UVP Bio-Lite) to moist beads on the column at 4 °C for 2.5 min followed by elution with chilled sodium phosphate buffer (100 µl). This was done a total of 4 times, and all elutions were combined with an extra 100 µl elution done at the end of the process, TCA-precipitated, and dissolved in a suitable volume of SDS dyes. The remainder of the sample left on the beads was eluted on the column by adding a suitable volume of SDS dyes and heating at 90 °C for 5 min. SDS PAGE was used for analysis.

2.1.1.2.1 Affinity Purification of Intact GGA PC Bio Tail Complex

Beads (20 µl of slurry) were prepared as described in Section 2.1.1.2. Then 10 µl of approximately 1 µM (less due to gel purification losses) 5' radiolabelled GGA PC Bio

METHODS

was pre-bound to the beads for 1 h at 4 °C before addition of the remaining sample ingredients to a final volume of 45 µl with final concentrations of 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, 40% pre-cleared (13000 RPM, 1 min) NE. The sample was incubated on ice for 20 min with periodic mixing, then placed onto column. The sample was washed and eluted as described in Section 2.1.1.2 except that elution was done with 25 mM Tris base, 200 mM glycine buffer and an extra 100 µl final wash was omitted. Control samples were set up as for native polyacrylamide gel electrophoresis and diluted (25 mM Tris base, 200 mM glycine) as necessary to give comparable radioactivity to the eluted sample. The remainder of the elution was separated in half and concentrated with either a 3 kDa or 10 kDa Microcon microconcentrator (Millipore) at 13000 RPM, 4 °C before electrophoresis. An equal volume of native gel dyes was added to all samples before running on a pre-chilled 10% native polyacrylamide gel (25 mM Tris base, 200 mM glycine, 10% glycerol) at 4 °C for 1 h at 150 V. The gel was then fixed (10% acetic acid, 10% IMS), dried and exposed to a phosphorimaging screen.

2.1.1.2.2 Affinity Purification of β -globin complex using GGA PC Bio

NeutrAvidin® beads were pre-washed as described in Section 2.1.1.2. Radiolabelled β -globin RNA (5 µl), which was modified with the addition of an annealing site for the photocleavable tail of GGA PC bio at its 3' end (Section 8.1.4), was pre-annealed to GGA PC Bio (1.5 µl of 5 µM) in a final volume of 10 µl by heating at 80 °C for 1 min, 30 °C for 10 min, then placing on ice for 2 min. Nuclear extract and remaining splicing components were pre-incubated with an oligonucleotide against the U6 snRNP (Dönmez *et al.*, 2007) in a final volume of 20 µl for 30 min at 30 °C to cause stalling at

METHODS

splicing complex A. The samples were then combined, giving a final volume of 30 μ l and final concentrations of 250 nM GGA PC Bio, 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, 1 μ M α -U6 oligo, and 40% commercial HeLa nuclear extract (Cilbiotech), and incubated a further 30 min at 30 °C to allow for the formation of splicing complex A. The sample was then bound to beads and washed as described in Section 2.1.1.2. Release of the complex from the beads was achieved by application of 365 nm UV light (UVP Bio-Lite) for 5 min followed by a 100 μ l elution (50 mM Tris base, 50 mM glycine). This was repeated and followed by an additional 100 μ l elution. All three elutions and remaining spin-through from column (~400 μ l) were combined and concentrated to ~ 40 μ l using a 10 kDa Microcon microconcentrator (Millipore) at 13000 RPM and 4 °C for 38 min. The sample was separated into halves, and one half was treated with 0.2 mg/ml heparin. Control samples for the native gel analysis were chilled and diluted (~150x in 50 mM Tris base, 50 mM glycine) after treatment with 0.8 mg/ml heparin until a sample of 20 μ l was comparable in radioactivity to the purified sample. Then 5 μ l of glycerol with dyes (bromophenol blue, xylene cyanol) was added and samples were loaded on a pre-chilled 2% native low melting point (LMP) agarose gel and run at 100 V for 4 hrs at 4 °C. The gel was then compressed, dried and exposed to a phosphorimaging screen.

2.1.2 Trichloroacetic Acid Precipitation of Proteins

Samples were chilled on ice for 30 min after addition of trichloroacetic acid to a final concentration of 20%. After centrifugation at 13,000 RPM for 10 min, the pellets were washed with 100 μ l chilled ethanol and dried. Samples were dissolved in a suitable volume of SDS dyes.

METHODS

2.1.3 SDS PAGE

All SDS PAGE gels consisted of a resolving gel (370 mM Tris Base pH 8.6, 12% acrylamide (ProtoGel, National Diagnostics), 0.1% SDS polymerized with 0.1% AMPS, 0.0013% TEMED) with a small amount of stacking gel on top (125 mM Tris Base pH 6.8, 4% acrylamide, 0.1% SDS polymerized with 0.1% AMPS 0.002% TEMED). Gels were run in SDS PAGE running buffer (125 mM Tris Base, 0.96 M Glycine, 0.1% SDS, pH 8.3) at 80 V through the stacking gel, then at 150 V for the remainder of the run time. Precision Plus ProteinTM Kaleidoscope protein marker (Bio-Rad) was used to track the progression of the samples through the gel. Gels were then stained or transferred to nitrocellulose as required.

2.1.3.1 Silver Staining

SDS PAGE gels were soaked for at least 5 min, at most overnight, in 50 ml of a 50% acetone solution with 0.37% formaldehyde and 0.04% trichloroacetic acid. The gels were rinsed 3 times with DI water, washed for 5 min in 50 ml water, then rinsed 3 more times, soaked in 50 ml of 50% acetone solution for 5 minutes, rinsed 3 times in water, then soaked for 1 min in 50 ml of 0.011% Na₂S₂O₃ solution. After 3 more water rinses, gels were soaked for 8 min in 50 ml of 0.28% AgNO₃ (w/v) with 0.74% formaldehyde. After rinsing 3 times in water, gels were developed with 50 ml of a solution containing 2% Na₂CO₃ (w/v), 0.0027% Na₂S₂O₃ and 0.016% formaldehyde and quenched by removing the developing solution and quickly adding 50 ml of a 50% acetone solution containing 1% acetic acid.

METHODS

2.1.3.2 Semi-dry Transfer of SDS PAGE Gel to Nitrocellulose

After electrophoresis, gels were soaked briefly in approximately 50 ml of transfer buffer (48 mM Tris base, 39 mM glycine, 0.0375% SDS, 10% methanol) then placed onto nitrocellulose (Amersham™ Hybond™ - ECL, GE Healthcare) on a stack of 8 pieces of 3MM chromatography paper (Whatman), damp with transfer buffer. Transfer buffer was added at each stage of assembly to ensure sufficient moisture and to prevent bubbles. Another similar stack of paper was placed on top of the gel, and transfer was done by running (Biometra Fast-Blot B33) at 10 W for 30 min.

2.1.4 Western Blot Analysis

Samples were run on an SDS PAGE gel and transferred to nitrocellulose. The nitrocellulose membrane was blocked overnight at 4 °C in blocking buffer consisting of 5% milk, 0.1% Tween-20, 1xTBS (20 mM Tris-HCL pH 7.5, 150 mM NaCl). It was then incubated for 1 h at 4 °C in the relevant dilution of primary antibody in blocking buffer. After washing the membrane 3 times for 5-10 min with 10 ml blocking buffer, the secondary antibody (1:1000 dilution in blocking buffer of Protein A/G HRP conjugate) was added and allowed to bind for 1 h at 4 °C. The blot was then washed three more times, as previously, before developing with ECL reagents (Pierce) according to the protocol provided. X-ray film (Fuji) was used to detect the ECL reaction. It is important to note that T7 tag western blots did not require a secondary antibody step as the HRP was conjugated directly to the anti-T7 antibody (Pierce).

METHODS

2.1.5 UV Crosslinking with Oligonucleotides

Samples (5 or 10 μ l) containing 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, 40% nuclear extract, and the desired amount of 5' or 3' radiolabelled oligonucleotide, were set up and incubated at 30 °C for 10 min, or as long as indicated for time-courses. Samples were then moved to a microtitre plate, if necessary, and 254 nm UV light (UVP SpotCure) was applied for 5 min. An equal volume of SDS dyes was added and samples were heated to 80 °C for 30 seconds before running on an SDS PAGE gel. Gels were then transferred to nitrocellulose or fixed (10% IMS, 10% acetic acid, 20 min) and dried before exposure to a phosphorimaging screen.

2.1.5.1 UV Crosslinking with GGA during SMN2 Splicing

Initially, half of the final reaction volume was set up containing 150 nM 3' radiolabelled GGA, 125 nM cold SMN2, 100 mM K/glu, 40 mM Hepes pH 7.5, and annealing was achieved by incubation at 30 °C for 30 min. Then the CrPi, MgCl₂, rATP, and commercial HeLa nuclear extract (Cilbiotech) were added, giving the final concentrations described in Section 2.1.5. The sample was incubated at 30 °C, and 5 μ l aliquots were taken into a microtitre plate for 254 nm UV application at various timepoints. Sample analysis was subsequently carried out as described in Section 2.1.5.

2.1.5.2 UV Crosslinking with G-quadruplex Stabilizing Ligands

G-quadruplex stabilizing ligand (50 μ M GQC-05, GSA-0820, or GSA-0902 in 10% DMSO) was pre-incubated with an equal volume of the desired 5' radiolabelled oligonucleotide for 30 min at 30 °C (1 μ l final volume). Simultaneously, the remaining

METHODS

ingredients, as described in Section 2.1.5, were pre-incubated at 30 °C for 30 min (9 µl final volume). Samples were then combined, giving the final concentrations mentioned in Section 2.1.5, with the addition of 2.5 µM ligand, and incubated at 30 °C for a further 10 min before 254 nm UV application. Sample analysis was subsequently carried out as described in Section 2.1.5. All handling was done in silanized plasticware, and commercial HeLa nuclear extract (Cilbiotech) was used.

2.2 *In vitro* RNA Techniques

2.2.1 Purification Techniques

2.2.1.1 Purification of Oligonucleotides using DNA/RNA Shadowing

Typically around 500 pmoles of the DNA or RNA to be purified was mixed 50% v/v with formamide dyes, heated to 80 °C for 1 min and run on a denaturing polyacrylamide gel (6% to 15% acrylamide, depending on the size of oligonucleotide and the resolution needed). The gel was transferred from the glass plates into Saran wrap and placed on top of a piece of thin layer chromatography (TLC) paper. The band(s) of DNA/RNA were visualized for excision by using a wand to apply long wave UV light, which caused the DNA/RNA to appear as a purple shadow against the TLC paper. Elution from the gel was carried out overnight in 350 µl sterile deionized water at 4 °C. The gel was then removed and the sample was precipitated with ethanol and dissolved in a suitable volume of TE.1. The concentration of the purified DNA/RNA was determined using a Nanodrop. Samples were stored at -20 °C or -80 °C as appropriate.

METHODS

2.2.1.2 Gel Purification of Radiolabelled RNA and Oligonucleotides

Samples were heated to 80 °C for 30 sec – 1 min and run on a 6% denaturing polyacrylamide gel. Purification from the gel was done using x-ray film (Fuji) and phosphorescent stickers to locate the desired band for excision. RNA was eluted out of the gel slice overnight at 4 °C in 350 µl elution buffer (500 mM Na acetate pH 5.2, 0.2% SDS, 1 mM EDTA pH 8), for large transcripts, or in 350 µl sterile deionized water for small oligonucleotides. The gel slice was then removed and the RNA was precipitated with ethanol, dissolved in the desired quantity of sterile deionized water and stored at -80 °C.

2.2.1.3 Proteinase K Treatment

Samples were treated with 0.4 mg/ml proteinase K (Roche) in proteinase K buffer (100 mM Tris-HCl pH 7.5, 12.5 mM 0.5 M EDTA pH 8.2, 150 mM NaCl, 1% SDS) at 37 °C for 15 min. To purify splicing reaction samples, 50 µl of this solution was used.

2.2.1.4 Precipitation with Ethanol

Unless a sample already contained a suitable salt content, sodium acetate was added to a final concentration of 300 mM. Then 2-3 volumes of ethanol were added, as appropriate. Typically samples were kept at room temperature during centrifugation (10-15 min). However, small oligonucleotides were pre-chilled by placement at -80 °C for 30 min prior to centrifugation at 4 °C (30 min). After initial centrifugation, pellet was washed with a suitable volume of ethanol (typically 200 µl, chilled if using for small oligonucleotides), then centrifuged again for 5-10 min before drying.

METHODS

2.2.2 Oligonucleotide Techniques

2.2.2.1 Analysis of Complex Formation on Oligonucleotides using Native Polyacrylamide Gel Electrophoresis

Samples of 5 or 10 μ l were set up containing 20 mM CrPi, 3.2 mM MgCl₂, 20 mM HEPES pH 7.5, 1.53 mM rATP, 50 mM K/glu, 40% nuclear extract, and the necessary amount of 5' radiolabelled oligonucleotide, and incubated at 4 °C for 30 min, or at 30 °C for 10 min, or as indicated. An equal volume of native gel dyes (25 mM Tris base, 200 mM glycine, 40% glycerol, bromophenol blue, xylene cyanol) was added before running on a pre-chilled native polyacrylamide gel (6-10% acrylamide, 25 mM Tris base, 200 mM glycine, with 10% glycerol in some cases) at 4 °C. Gels were fixed (10% acetic acid, 10% IMS, 20 min), as necessary, and dried before exposure to a phosphorimaging screen. Note that the native gels analyzing the effects of G-quadruplex stabilizing ligands (GQC-05, GSA-0820, GSA-0902) were set up exactly as for UV crosslinking (1/2 of the same exact samples in fact).

2.2.2.2 5' Radiolabelling of DNA/RNA Oligonucleotides

Equimolar amounts of the oligonucleotide to be labeled and [γ -³²P] rATP (10 mCi/ml, 3000 Ci/mmol) (Perkin Elmer) were incubated in buffer (50 mM Tris pH 7.5, 10 mM MgCl₂, 1 mM DTT) with 5% T4 polynucleotide kinase (NEB) in a final volume of 10 μ l for 30 min at 37 °C. Then the enzyme was deactivated by heating at 70 °C for 10 min. When necessary, unreacted [γ -³²P] rATP was removed using either a G25 spin column (GE Healthcare) or denaturing polyacrylamide gel purification.

METHODS

2.2.2.3 3' Radiolabelling of DNA/RNA using [5'-³²P]pCp (cytidine-3',5'-bis-phosphate)

Initially, 12 pmol of radioactive pCp was made by incubating 12 pmol of [γ -³²P] rATP (10 mCi/ml, 3000 Ci/mmol) (Perkin Elmer) and 12 pmol of 3' cytidine monophosphate in buffer (50 mM Tris pH 7.5, 10 mM MgCl₂, 1 mM DTT) with 5% T4 polynucleotide kinase (10 units/ μ l, NEB) in a final volume of 10 μ l at 37 °C for 30 min. The enzyme was deactivated by heating at 70 °C for 10 min. Then the 3' radiolabelling was done overnight at 4 °C by adding to the chilled pCp sample: 10 pmol of the DNA/RNA to be labeled, Hepes pH 7.5 (37.5 mM final concentration), MgCl₂ (15 mM), DTT (2.25 mM), DMSO (10%) and 2 units of T4 RNA ligase (10 units/ μ l, Promega), giving a final volume of 20 μ l. The sample was then gel purified.

2.2.2.4 Reverse Transcription on Oligonucleotides

Initially, samples were annealed by slow cooling from 80 °C to room temperature in a volume of 5 μ l containing final concentrations of 40 nM 5' radiolabelled GGA DNA primer, 400 nM bifunctional oligonucleotide of interest and 1 x M-MuLV Reverse Transcriptase Reaction Buffer (NEB). Samples were placed on ice for 10 min, then moved to room temperature before addition of 0.2 μ l 5 mM dNTPs and 0.2 μ l M-MuLV reverse transcriptase (200 units/ μ l, NEB M0253). Samples were left at room temperature for 20 min. Then samples were rapidly placed at 80 °C for 1-2 min after addition of 5 μ l of formamide dyes. Samples were run on a 15% denaturing polyacrylamide gel at 36 W, then fixed with a solution of 10% acetic acid, 10% IMS for 20 min before drying. Radioactivity was detected by exposure to a phosphorimaging

METHODS

screen, which was read by a Cyclone or Typhoon™ phosphorimager and visualized using OptiQuant™ software.

2.2.2.5 Preparation of Small Molecular Weight Radioactive Marker

The small molecular weight marker was generated by digestion of pBR322 plasmid (1 µg, Fermentas) with HaeIII restriction enzyme (5 units, NEB) for 30 min at 37 °C in 1 x Buffer 4 (NEB) in a final volume of 5 µl. The sample was then made up to 10 µl final volume for treatment with antarctic phosphatase (5 units, NEB) in 1 x antarctic phosphatase buffer (NEB) at 37 °C for 15 min. Enzymes were deactivated by heating at 65 °C for 15 min. Then 1 µl of sample was 5' radiolabelled by incubating with 1 µl [γ -³²P] rATP (10 mCi/ml, 3000 Ci/mmol) (Perkin Elmer) in buffer (50 mM Tris pH 7.5, 10 mM MgCl₂, 1 mM DTT) with 5% T4 polynucleotide kinase (10 units/µl, NEB) for 30 min at 37 °C. The enzyme was deactivated by heating at 70 °C for 10 min. Samples were diluted with formamide dyes, as necessary, and stored at -80 °C.

2.2.3 Glycerol Gradients

2.2.3.1 Glycerol Gradients with Radiolabelled GGA

Initially, 7 µl samples containing 86 nM 3' radiolabelled GGA, 107 nM K⁺/glu, and 43 nM HEPES, pH 7.5 were set up with and without 107 nM cold SMN2, and annealing was allowed to progress at 30 °C for 30 min. Meanwhile, the nuclear extract was pre-incubated with an oligonucleotide against the U6 snRNP (Dönmez *et al.*, 2007), CrPi, MgCl₂, and rATP for 30 min at 30 °C. Then 8 µl of the nuclear extract mix was combined with the 7 µl annealed sample, giving final concentrations of 20 nM GGA, (50 nM

METHODS

SMN2), 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, 1 μM α-U6 oligo, and 40% commercial HeLa nuclear extract (Cilbiotech), and incubated for a further 30 min at 30 °C to allow for the formation of splicing complex A. Samples were then UV crosslinked, chilled, and loaded onto a glycerol gradient (10 – 50% glycerol, 20 mM Hepes pH 8, 0.2 mM EDTA, 1 mM DTT, 100 mM KCl) (Gradient Master ip 107, BioComp). The gradient was centrifuged for 16 ½ hours at 4 °C using a Sorval rotor (660) at 23400 RPM. Gradients were fractionated on ice (175 μl fractions), and the radioactivity in each fraction was measured using Cherenkov scintillation counting. Protein was precipitated with TCA and analyzed by SDS PAGE.

2.2.3.2 Glycerol Gradient with Radiolabelled β-Globin RNA

Initially, nuclear extract was pre-incubated with an oligonucleotide against the U6 snRNP (Dönmez *et al.*, 2007), CrPi, MgCl₂, and rATP for 30 min at 30 °C in a final volume of 8 μl. Then a 7 μl sample of radiolabelled β-globin RNA, K/glu and Hepes pH 7.5 was added, giving final concentrations of 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, 1 μM α-U6 oligo, and 40% commercial HeLa nuclear extract (Cilbiotech). The sample was incubated a further 30 min at 30 °C, before being chilled and loaded onto a glycerol gradient (10 – 50% glycerol, 20 mM Hepes pH 8, 0.2 mM EDTA, 1 mM DTT, 100 mM KCl)(Gradient Master ip 107, BioComp). The gradient was centrifuged for 16 ½ hours at 4 °C using a Sorval rotor (660) at 23400 RPM. It was then fractionated on ice (175 μl fractions), and the radioactivity in each fraction was measured using Cherenkov scintillation counting.

METHODS

2.2.4 *In vitro* Transcription

2.2.4.1 *In vitro* Transcription of RNA without Radiolabelling ('Cold' Transcription)

A 100 µl transcription reaction containing 40 mM Tris pH 7.5, 20 mM MgCl₂, 10 mM NaCl, 2 mM spermidine HCl, 10 mM DTT, 4 mM rNTPs, 10 ng/µl linear DNA template (prepared by PCR), 5% RNase OUT (Invitrogen), 5% T7 polymerase (1:20) was incubated at 37 °C for 4 hrs. To stop the transcription, the reaction volume was doubled with the addition of 10x DNase buffer (Promega) to a final concentration of 1x, DNase (Promega, 1 unit/µl) to a final concentration of 10% v/v, and water. The sample was incubated for 30 min at 37 °C, then quenched with the addition of 10% of the reaction volume of stop solution (Promega) and incubation at 65 °C for 10 min. DNA fragments were removed using an S-300 column (GE Healthcare), and the remaining material was extracted with phenol-chloroform and precipitated using ethanol. The sample was dissolved in 20% of the original reaction volume of TE.1. The concentration of RNA was determined using a Nanodrop, and sample quality was checked by polyacrylamide gel. RNA was stored at -80 °C.

2.2.4.2 *In vitro* Transcription of ³²P Radiolabelled RNA ('Hot' Transcription)

A 10 µl transcription reaction containing 40 mM Tris HCl pH 7.5, 6 mM MgCl₂, 2 mM spermidine HCl, 10 mM NaCl, 0.5 mM rATP, 0.5 mM rCTP, 0.5 mM rUTP, 0.05 mM rGTP, 1 mM Diguanosine Triphosphate Sodium [G(5')ppp(5')G] (GE Healthcare), 5 mM DTT, 5 ng/µl linear DNA template (prepared by PCR), 5% RNase OUT (Invitrogen), 0.33

METHODS

μM [α - ^{32}P] rGTP (10 mCi/ml, 3000 Ci/mmol) (Perkin Elmer), 5% T7 polymerase (1:20) was set up at room temperature, then incubated at 37 °C for 1-2 hrs. The transcript was purified in a denaturing gel, recovered by elution and precipitation, and dissolved in about double the original reaction volume of sterile purified water. Sample was stored at -80 °C.

2.2.5 *In vitro* Splicing

Samples of 10 or 15 μl containing a suitable amount of radiolabelled RNA transcript, 20 mM CrPi, 3.2 mM MgCl_2 , 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, and 40% nuclear extract were set up at 4 °C then incubated at 30 °C for 2 hrs. Aliquots of 1.5 or 2 μl were taken into a microtitre plate on dry ice at designated timepoints, typically 0, 30, 60, and 120 min. The plate was then thawed and samples were proteinase K (PK)-treated, ethanol-precipitated to purify the RNA, and dissolved in 10 μl formamide dyes. The samples were heated at 80 °C for 30 sec, then run on a 6% denaturing polyacrylamide gel, which was subsequently dried and exposed to a phosphorimaging screen. Quantification of pre-mRNA and both mRNA products was done using OptiQuant software, and intensities were adjusted to account for the number of radioactive guanosines in each molecular species.

2.2.5.1 *In vitro* Splicing with Oligonucleotides

Oligonucleotide at the desired concentration was pre-annealed to the radiolabelled RNA transcript in 100 mM K/glu, 40 mM Hepes pH 7.5 by slow cooling in a PCR machine (80 °C 5 min, 70 °C 5 min, 60 °C 5 min, 50 °C 5 min, 40 °C 5 min, 30 °C 5 min). One annealing mix was made for all similar samples in order to give a decent volume

METHODS

for the slow cooling procedure. This mix was then aliquotted (2.5 or 5 μ l) and an equal volume of a mix of remaining ingredients was added on ice, giving a final volume of 5 or 10 μ l and final concentrations of 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, and 40% nuclear extract. Samples were then incubated and analyzed as described in Section 2.2.5.

2.2.5.2 *In vitro* SMN2 Splicing with GGA PC Bio Cleavage

GGA PC Bio (500 nM) was pre-annealed to the radiolabelled SMN2 transcript in a final volume of 22.5 μ l in 100 mM K/glu, 40 mM Hepes pH 7.5 at 30 °C for 10 min. A mock annealing lacking GGA PC bio was also done. Then an equal volume of a mix of remaining ingredients was added on ice, giving a final volume of 45 μ l and final concentrations of 250 nM GGA PC Bio (not in control sample), 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, and 40% commercial HeLa nuclear extract (Cilbiotech). Samples were separated into 5 μ l aliquots in 1.5 ml tubes for each cleavage timepoint and incubated at 30 °C for 2 hrs in the dark. Cleavage of GGA PC Bio was achieved by application of 365 nm UV light from a UVP Bio-Lite machine for 5 minutes in a 30 °C block at the designated timepoints. Samples were then processed as described in Section 2.2.5.

2.2.5.3 *In vitro* SMN2 Splicing with G-quadruplex Stabilizing Ligands

GGA (500 nM) in the presence of the desired concentration of G-quadruplex stabilizing ligand (GQC-05, GSA-0820, GSA-0902 in 10% DMSO) was pre-annealed to the radiolabelled SMN2 transcript in 100 mM K/glu, 40 mM Hepes pH 7.5 at 30 °C for 10 min. One annealing mix was made for all similar samples. Then 2.5 μ l aliquots were

METHODS

taken and an equal volume of a mix of remaining ingredients was added on ice, giving a final reaction volume of 5 μ l with final concentrations of 250 nM GGA, 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, 40% commercial HeLa nuclear extract (Cilbiotech), and either 25 nM, 250 nM or 2.5 μ M ligand (final concentration of DMSO was 0.5%). Samples were then incubated at 30 °C for 2 hrs and processed as described in Section 2.2.5. All handling was done in silanized plasticware. G-quadruplex stabilizing ligands were kindly provided by Professor Laurence Hurley, University of Arizona.

2.2.5.4 *In vitro* SMN2 Splicing with Purified hnRNP F Domains

GGA was pre-annealed to radiolabelled SMN2 in 167 mM K/glu, 67 mM Hepes pH 7.5 by slow cooling in a PCR machine (80 °C 5 min, 70 °C 5 min, 60 °C 5 min, 50 °C 5 min, 40 °C 5 min, 30 °C 5 min). One annealing mix was made for all similar samples in order to give a decent volume for the slow cooling procedure. Splicing samples were set up by combining 1 μ l of a suitable concentration of purified hnRNP F domain(s) (Dominguez & Allain, 2006) in the designated buffer, 2.5 μ l nuclear extract mix (containing CrPi, MgCl₂ and rATP), and 1.5 μ l pre-annealed SMN2/GGA in that specific order, giving final concentrations of 250 nM GGA, 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, 40% commercial HeLa nuclear extract (Cilbiotech). Samples were incubated at 30 °C for 2 hrs and processed as described in Section 2.2.5.

2.2.5.5 *In vitro* Splicing with Gold Nanoparticles

Initially, the desired concentration of gold nanoparticle or control oligonucleotide was pre-incubated with radiolabelled RNA transcript in a final volume of 1.5 μ l at 30 °C for

METHODS

10 min to allow for annealing. Then 3.5 μ l of a mix of remaining splicing ingredients was added on ice, giving final concentrations of 20 mM CrPi, 2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, and 40% commercial HeLa nuclear extract (Cilbiotech), which was previously dialyzed to remove DTT (10% glycerol, 20 mM Hepes pH 8, 0.2 mM EDTA, 100 mM KCl). Samples were incubated at 30 °C for 2 hrs and processed as described in Section 2.2.5. However, samples were not heated prior to gel electrophoresis.

2.2.5.6 *In vitro* SMN2 Splicing with Morpholino Oligonucleotides

The desired concentration of morpholino oligonucleotide (PMO-18, PMO-20, PMO-25 (Zhou *et al.*, 2013)) was pre-annealed to the radiolabelled SMN2 in 2/5 of the final reaction volume by incubation at 30 °C for 10 min in silanized tubes. A mix of the remaining splicing components was added, giving a final volume of 5 μ l and final concentrations of 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, and 40% commercial HeLa nuclear extract (Cilbiotech). Samples were incubated at 30 °C for 2 hrs and processed as described in Section 2.2.5.

2.2.5.7 Preparation of Normal Radioactive Marker

The marker was generated by digestion of pBR322 plasmid (2 μ g, Fermentas) with HPA II restriction enzyme (10 units, NEB) for 30 min at 37 °C in 1 x Buffer 1 (NEB) in a final volume of 10 μ l. Sample was then made up to 20 μ l final volume for treatment with antarctic phosphatase (10 units, NEB) in 1 x antarctic phosphatase buffer (NEB) at 37 °C for 15 min. Enzymes were deactivated by heating at 65 °C for 15 min. Then 1 μ l of sample was 5' radiolabelled by incubating with varying amounts (0.2 μ l - 1 μ l,

METHODS

depending on level of radioactivity desired) of [γ - 32 P] rATP (10 mCi/ml, 3000 Ci/mmol) (Perkin Elmer) in buffer (50 mM Tris pH 7.5, 10 mM MgCl₂, 1 mM DTT) with 5% T4 polynucleotide kinase (10 units/ μ l, NEB) for 30 min at 37 °C. The enzyme was deactivated by heating at 70 °C for 10 min. The sample was diluted with formamide dyes, as necessary, and stored at -80 °C.

2.2.6 Analysis of Splicing Complexes A, B and C by Native

Agarose Gel Electrophoresis

Samples were set up and timepoints taken as described in Section 2.2.5. Samples were then treated with 0.8 mg/ml heparin for 30 mins at room temperature. An equal volume of loading dyes (50 mM tris, 50 mM glycine, 40% glycerol, xylene cyanol, bromophenol blue) was added before running samples on a 2% native LMP agarose gel (50 mM Tris base, 50 mM glycine) at 100 V for 4 hrs at 4 °C (Das & Reed, 1999). The gel was then compressed overnight between two sheets of 3MM chromatography paper (Whatman) surrounded by paper towels, in order to remove excess liquid and make it thin enough to be dried and exposed to a phosphorimaging screen. To stall the splicing reaction at A complex, an oligonucleotide against the U6 snRNP (Dönmez *et al.*, 2007) was pre-incubated with the nuclear extract in the presence of ATP for 30 min at 30 °C, giving a final concentration of 1 μ M in the final splicing reaction.

2.2.7 Quantification of Gels with Radioactive RNA/DNA

The radiation from various types of dried gels was detected using a storage phosphor imaging system. The gels were exposed to a phosphorimaging screen for a suitable length of time dependent upon the level of radioactivity. Then the screens were read

METHODS

by a Cyclone or Typhoon™ phosphorimager and visualized using OptiQuant™ software. This software was also used to quantify the radioactive intensity of certain gel bands when necessary. A background baseline level of radioactivity was subtracted from the intensity of any given band prior to analysis. Also, for splicing gels, intensities were adjusted to account for the number of radioactive guanosines in each molecular species.

2.3 Cell Culture and Nuclear Extract Preparation

2.3.1 Transfection of HEK 293T Cells

Cells were seeded onto 15 cm² plates (Corning) with a density of 5 x 10⁴ cells per plate, as determined using a haemocytometer with 1 mm² grids. Cells were allowed to grow to 50% confluency in 20 ml complete media (DMEM (Gibco®), 10% FBS (Gibco®), 1% PenStrep (Gibco®)). Then the medium was changed 4 hrs before beginning calcium chloride transfection. The desired amount of plasmid (12-24 µg) in TE.1 was mixed on ice with 10 x the volume of CaCl₂ solution (1 mM Tris-HCl pH 7.5, 0.1 mM EDTA, 300 mM CaCl₂). An equal volume of pre-chilled HBS solution (342.23 mM NaCl, 12.41 mM KCl, 1.76 mM Na₂HPO₄ (anhydrous), 13.88 mM glucose, 52.45 mM Hepes pH 7.5) was added and incubated on ice for 10 minutes. The mix (3.5 ml) was carefully added dropwise to the plate before 16-24 h incubation. Cells were then shocked for 3 min with 20 ml DMSO-DMEM solution (75% DMEM, 25% DMSO), washed twice with 8 ml complete medium, and then incubated for 16-24 h in 20 ml fresh complete medium.

METHODS

2.3.2 Nuclear Extract Preparation

After transfection, cells were harvested from the plates in 10 ml chilled PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.4), pelleted, washed with 1 ml PBS and collected by centrifugation. The packed cell volume (PCV) was estimated, and cells were resuspended in one PCV of Buffer A (10 mM Hepes pH 8, 1.5 mM MgCl₂, 10 mM KCl, 1 mM DTT, 0.06% NP40). After swelling on ice for 15 min, cells were lysed by vortexing vigorously for 10 seconds. Nuclei were pelleted by centrifugation for 30 sec at 13000 RPM at 4 °C, then resuspended in 0.7 x PCV Buffer C (20 mM Hepes pH 8, 25% glycerol, 420 mM NaCl, 0.2 mM EDTA, 1 mM DTT). Sample was incubated for 30 min at 4 °C with mixing using small magnetic fleas, and then centrifuged at 13000 RPM for 5 min to remove debris. Nuclear extract was dialyzed against Buffer D (10% glycerol, 20 mM Hepes pH 8, 0.2 mM EDTA, 1 mM DTT, 100 mM KCl), snap-frozen with liquid nitrogen and stored at -80 °C.

2.3.3 Dialysis of Nuclear Extract

Nuclear extract was dialyzed against the desired buffer for 2 hrs at 4 °C in 50 µl drops on dialysis membranes (0.025 µm, Millipore) floating in petri dishes.

2.4 Single Molecule Techniques

2.4.1 Preparation of the Sample Chamber

Glass slides were sonicated for 10 min in purified water, dried by aspiration, and then cleaned further with 100% argon plasma at a pressure of 0.15 mbar at 80 W for 2 min (MiniFlecto-PC-MFC). A sample chamber was then constructed using double-sided tape

METHODS

to make a 5 mm wide channel down the middle of the slide (22 mm long, ~200 μm thick), which was sealed with a small glass coverslip. The binding surface was generated by application of a biotin-BSA solution (20 $\mu\text{g}/\text{ml}$ in PBS, Thermo Scientific) for 10 min, followed by a wash with SM buffer (100 mM NaCl, 50 mM Hepes pH 7.5, 20 units/ml RNase OUT (Invitrogen)), and then a 10 min incubation with a streptavidin solution (10 $\mu\text{g}/\text{ml}$ in PBS, Invitrogen). All solutions were pre-chilled at 4 $^{\circ}\text{C}$, but sample chamber was kept at room temperature. The surface was washed again with SM buffer to remove excess streptavidin.

2.4.2 Single Molecule Experiment

2.4.2.1 Labelling RNA with β -globin 5' Cy5 Oligonucleotide

SMN2 RNA at a concentration of 1 μM was labelled with Cy5 using an oligonucleotide complementary to the 5' end of the β -globin exon 2 in the SMN2 construct (β -globin 5' Cy5, Section 8.2.4). Various concentrations of β -globin 5' Cy5 were used, starting at a ratio of 1:1 with the SMN2 RNA and descending. The 5 μl annealings were done in 100 mM NaCl, 10 mM Hepes pH 8.0 by slow cooling in a hot block from 80 $^{\circ}\text{C}$ to 45 $^{\circ}\text{C}$, then placing on ice. Samples were analyzed by native polyacrylamide gel electrophoresis (1 x TBE, 6% acrylamide) to determine which sample(s) contained no free β -globin 5' Cy5. A Typhoon™ Imager was used to detect fluorescence originating from the Cy5 in the gel. No dyes were used in the samples, as they would cause interference with the imaging.

METHODS

2.4.2.2 Sample Preparation

Initially, 167 nM labelled SMN2 RNA was pre-incubated with 833 nM NT oligonucleotide (Section 8.2.1), or just water, for 30 min at 30 °C to allow for annealing. Meanwhile, the remaining splicing components, in a final volume of 3.5 µl, were pre-incubated at 30 °C for 30 min to allow α-U6 oligonucleotide to bind (Dönmez *et al.*, 2007). Then 1.5 µl of the RNA sample was added, giving final concentrations of 50% eGFP-Sam68 + T7-CNBP HEK 293T NE, 50 nM SMN2, (250 nM NT), 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, 1 µM α-U6 oligonucleotide and 3 units RNase OUT (Invitrogen). The sample was incubated a further 30 min at 30 °C to allow formation of splicing complex A, then chilled on ice prior to serial dilution with pre-chilled SM buffer (100 mM NaCl, 50 mM Hepes pH 7.5, 20 units/ml RNase OUT (Invitrogen)). First, the most dilute sample (1:6000) was placed into the sample chamber, which was mounted on the microscope, and allowed to bind for 5 min. The concentration of Cy5 and eGFP spots on the surface of the slide was briefly assessed, and more concentrated samples were added if necessary. Once a suitable sample concentration was obtained, imaging buffer (2 mM PCA, 0.1 µM PCD, 8 units (0.16 U/µl) RNase OUT (Invitrogen)) was added to the sample chamber and left for 5 min prior to data acquisition.

2.4.3 Data Acquisition and Analysis

A custom-built, objective-based total internal reflection (TIRF) microscope was used for imaging. Lasers of 488 nm, for detection of eGFP, and 640 nm, for detection of Cy5, with the same beam path were used. The lasers were set up in such a way as to achieve total internal reflection, resulting in limited illumination of approximately 100

METHODS

nm above the surface of the sample chamber. Therefore, only fluorophores in complexes attached to the surface were excited and detected. Control of the lasers and acquisition of the data were achieved using a LabVIEW program. Image frames of 250 x 250 pixels, with each pixel being approximately 160 nm wide, were obtained using an electron-multiplying charge-coupled device (EMCCD). Frames, each representing 100 milliseconds, were generated by integrating intensity over time. Initially, the 640 nm laser was used to detect the Cy5 labelled RNA for 50 frames. Then the 488 nm laser was used to detect the eGFP-labelled protein, and acquisition was continued until all proteins of interest were bleached (typically around 120 seconds). The data, stored as a multi-page TIFF file, was analyzed using a MATLAB program written by colleague Mr R. Weinmeister. The program generated a single time-averaged image for each of the wavelengths, showing each fluorophore as a visible spot. All Cy5-labelled RNA spots were counted, and compared with the overlaid eGFP labelled protein image to determine co-localization. The program identified spots as being co-localized if they were within 2 pixels of each other, and these results were confirmed by eye to be valid and logical. The number of eGFP bleaching steps for each co-localized spot was also assessed by the program and confirmed by eye, using background-corrected intensity traces. At least 100 co-localized spots were analyzed in this way for each experiment, giving a final assessment of percentage of co-localization between the Cy5-labelled RNA and the eGFP-labelled protein as well as a distribution of the number of eGFP-labelled proteins bound per RNA.

2.5 Techniques Related to Fe-BABE Experiments

2.5.1 Synthesis of Fe-BABE Modified Oligonucleotides

2.5.1.1 Synthesis of Thiol Modified Oligonucleotides

Oligonucleotides were synthesized using SynBase™ CPG cartridges, phosphoramidite bases, C6 S-S CE thiol-modifier phosphoramidite, and synthesis reagents from Link Technologies using automated solid phase synthesis.

2.5.1.2 Deprotection of Dithiol Protection Group on Oligonucleotides

The dithiol protection group on the modified oligonucleotides was removed by incubation of 800 pmol of oligo with 10 mM TCEP in 100 mM Tris pH 8.5 for 1 h at 30 °C. These conditions were confirmed using reverse phase HPLC. Excess TCEP was removed by dialyzing sample into sterile purified water for 30 min at room temperature, using a dialysis membrane (0.025 µm, Millipore).

2.5.1.3 Fe-BABE Conjugation

Initially, Fe-BABE was generated by incubation of 14.7 mM BABE (Dojindo) with 11.8 mM ammonium Fe(II) sulfate hexahydrate and 80 mM sodium acetate (not pHed) for 35 min at room temperature prior to use. A yellow color change was observed as the Fe was coordinated. The deprotected thiol-modified oligonucleotide sample (800 pmol) was then combined with 10 µl of the Fe-BABE mix and incubated for 1 h at 37 °C in the presence of 40 mM potassium phosphate pH 8.0. The sample was extracted with TE.1 saturated phenol to remove unreacted Fe-BABE then ethanol-precipitated. Samples were resuspended in sterile purified water at an estimated concentration of

METHODS

10 μ M, assuming a 50% loss throughout the procedure. More accurate determination of concentration using UV absorption was impossible due to interfering absorption of Fe-BABE at 260 nm.

2.5.2 Fe-EDTA experiments

Samples of 5 μ l volume were set up containing 50 nM 5' radiolabelled GGA-O (Section 8.2.1), 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, and 50% of either Buffer D (20 mM Hepes pH 8.0, 100 mM KCl, 10% glycerol, 0.2 mM EDTA, 1 mM DTT), Glycerol-free Buffer D (20 mM Hepes pH 8.0, 100 mM KCl, 0.2 mM EDTA, 1 mM DTT), or PEG Buffer D (20 mM Hepes pH 8.0, 100 mM KCl, 10% dimethyl PEG solution (70% w/v in H₂O, approx. 2000 MW), 0.2 mM EDTA, 1 mM DTT). Samples were incubated at 30 °C for 10 min prior to initialization of cleavage with 1/5 reaction volume of Fe-EDTA mix (equal volumes of 50 mM ammonium Fe(II) sulfate hexahydrate (Fe(NH₄)₂(SO₄)₂ 6H₂O), 100 mM EDTA, 250 mM ascorbate, 2.5% hydrogen peroxide (H₂O₂) added in the order indicated), which was mixed immediately prior to addition to the sample. After 10 min at 30 °C, the reaction was quenched by addition of thiourea to a concentration of 500 mM. Then an equal volume of formamide dyes was added and samples were heated at 80 °C for 2 min prior to running on a 20% denaturing polyacrylamide gel. The gel was fixed, dried, and exposed to a phosphorimaging screen.

METHODS

2.5.3 Fe-BABE experiments

2.5.3.1 Sample Preparation and Cleavage

Initially, 250 nM SMN2 was pre-annealed to ~500 nM Fe-BABE modified oligonucleotide in the presence of 100 mM K/glu and 40 mM Hepes pH 7.5 in a final volume of 20 μ l by slow cooling in a PCR machine (80 °C 5 min, 70 °C 5 min, 60 °C 5 min, 50 °C 5 min, 40 °C 5 min, 30 °C 5 min). Then 5 μ l H₂O was added and free Fe-BABE modified oligo (and buffer) was removed using an S300 spin column (GE Healthcare). Samples of 30 μ l were set up containing pre-annealed Fe-BABE oligo/SMN2 (~50 nM), 20 mM CrPi, 3.2 mM MgCl₂, 20 mM Hepes pH 7.5, 1.53 mM rATP, 50 mM K/glu, and 40% of either Glycerol-free Buffer D, commercial HeLa NE dialyzed in Glycerol-free Buffer D. When experiment involved the use of nuclear extract, 1 μ M α -U6 oligo was also added (Dönmez *et al.*, 2007). Samples were incubated at 30 °C for 20 min prior to initialization of free radical production, which was achieved by simultaneous mixing of the sample with 2 individual 0.6 μ l drops of 250 mM ascorbate and 2.5% H₂O₂ in a microtitre plate. Cleavage was allowed to progress at 30 °C for 10 min before reaction was quenched by the addition of 5 μ l of 80% glycerol. Samples were treated with PK, and the RNA was precipitated with ethanol after the addition of 1 μ l Pellet Paint® Co-precipitant (Millipore).

2.5.3.2 Reverse Transcription and Analysis

Ethanol precipitated samples were re-dissolved in 11 μ l H₂O and 2.25 μ l of this was pre-annealed to each of four 5' radiolabelled SMN2 primers (0.25 μ l of 1 μ M) by heating to 80 °C then placing directly onto ice. For control reactions, cold SMN2 at the

METHODS

same concentration as the primers was used. Samples were placed at room temperature prior to addition of an equal volume of reverse transcription master mix, giving final concentrations of 1 x reverse transcription buffer (Promega: 50mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM MgCl₂, 10 mM DTT), 0.25 mM dNTPs (and 0.25 mM ddNTP when indicated), 5 units RNase OUT (Invitrogen), and 38 units M-MLV Reverse Transcriptase, RNase H Minus (M530A, Promega). Reverse transcription was allowed to progress at 42 °C for 1 h. Then the enzyme was inactivated by incubation at 70 °C for 15 min. Condensation on the lid of the sample tube was collected by brief centrifugation. Then the RNA template was degraded by addition of sodium hydroxide to a final concentration of 200 mM followed by heating at 70 °C for 10 min. An equal volume of formamide dyes were added and samples were heated briefly at 80 °C prior to running on a 6% denaturing polyacrylamide gel. The gel was then dried, exposed to a phosphorimaging screen, and imaged using a Typhoon™ imager.

3 BIOTIN AFFINITY PURIFICATION WITH PHOTOCLEAVABLE LINKER

3.1 Introduction

3.2 Design of Oligonucleotide with Photocleavable Linker

3.3 Optimization of Method

3.4 Extension of Method to Purify Intact Complexes

3.5 Discussion

3.1 Introduction

The oligonucleotide of interest in this study, named GGA, was first developed in 2003 as a potential therapy for spinal muscular atrophy (Skordis *et al.*, 2003). It dramatically increases the level of SMN2 exon 7 inclusion during splicing by annealing to SMN2 exon 7 and recruiting positive splicing factors with its enhancer tail. The concept is relatively simple, but determining how exactly the enhancer tail works required further investigation. The most obvious starting point was to determine what proteins were binding. Commonly this has been done using affinity purification techniques.

Biotin affinity purification is a robust traditional purification method that exploits the very strong binding affinity of the small molecule biotin to avidin or similar proteins. There are several different ways to conjugate biotin to DNA, RNA or proteins of interest, making the method quite versatile as well. The limitation of this method arises as a result of the strong binding affinity, the quality that also makes it such a powerful tool. This is because the only way to harvest the sample at the end of the purification procedure is through the use of harsh conditions, usually elevated temperature. These conditions cause all background proteins bound to the resin to be released into the sample of interest and prevent further analysis of the sample in a biologically relevant state. These problems might be avoided by using a biotinylated oligonucleotide with a photocleavable linker, which can be easily cleaved using 365 nm UV light (UVP Bio-Lite) under gentle, biologically relevant conditions. This chapter describes experiments done to test this approach.

3.2 Design of Oligonucleotide with Photocleavable Linker

As mentioned in the introduction, the photocleavable linker technique has the potential to reduce background and potentially allow for elution of intact complexes. In addition, the use of the photocleavable linker in this situation allowed for the enhancer tail to be eluted separately from the annealing region, as shown in Figure 3.1. This modified version of GGA was named GGA PC Bio (Section 8.2.1).

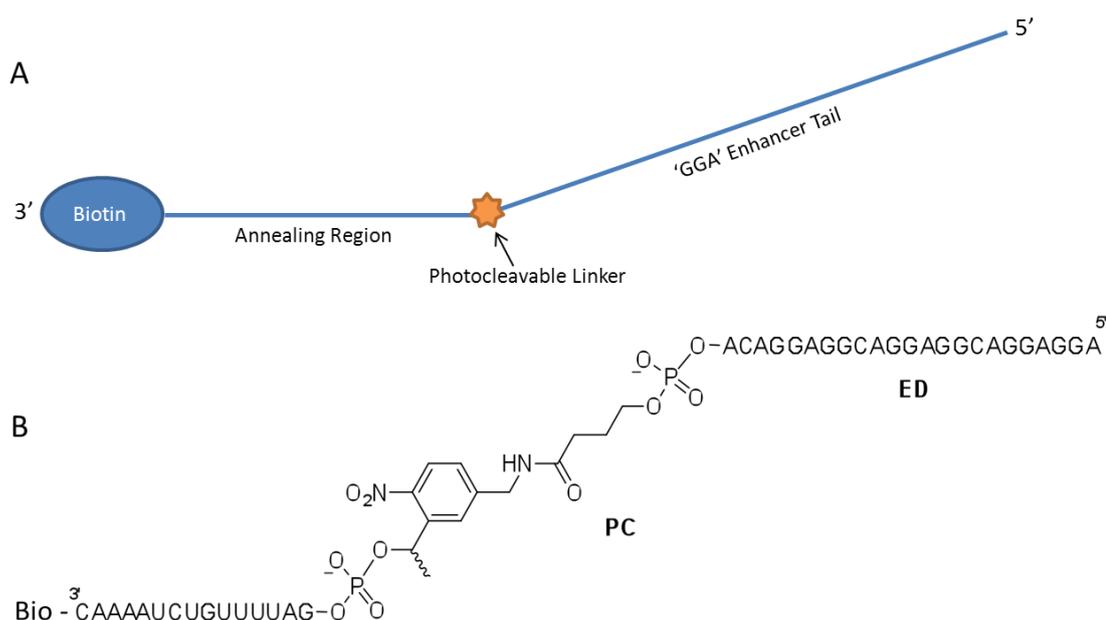


Figure 3.1: Design of GGA PC Bio, oligonucleotide for biotin affinity purification with photocleavable linker.

(A) Cartoon diagram of GGA PC Bio. **(B)** Diagram of GGA PC Bio including sequence information and photocleavable linker chemistry (Smith *et al.*, manuscript submitted)

3.3 Optimization of Method

Before optimization of the photocleavage technique was embarked upon, the conditions for traditional biotin affinity purification were optimized using a biotinylated 2'OMe oligonucleotide designed to anneal to the U1 snRNP (α -U1 Bio, Section 8.2.4). This was chosen as it was readily available and the target proteins to be

purified were already known (Seiwert & Steitz, 1993). Many tests of buffer and washing conditions were done. However, the only condition found to be extremely important for the reduction of background protein binding was the presence of detergent during all steps involving the NeutrAvidin® beads (Figure 3.2).

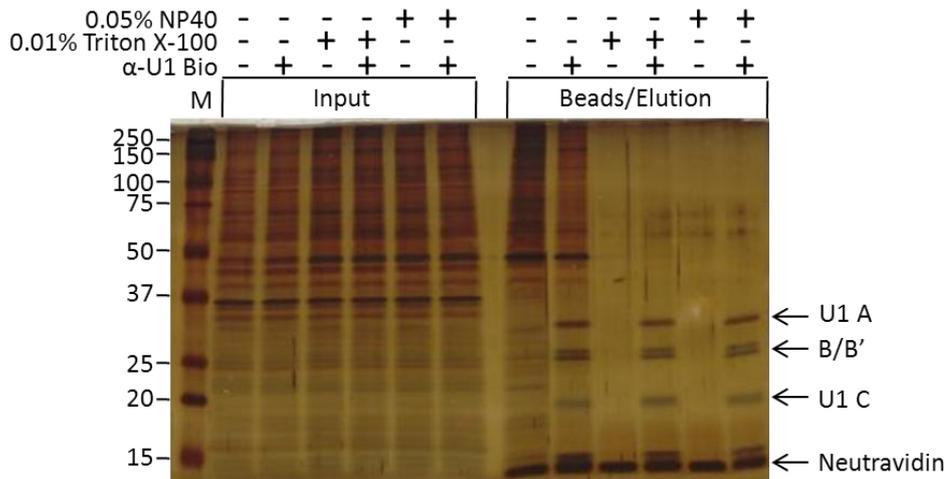


Figure 3.2: Detergent is vital for clean affinity purifications.

This silver stained SDS PAGE gel of test affinity purifications from commercial HeLa nuclear extract using 250 nM α-U1 Bio, a biotinylated 2'OMe oligo targeting the U1 snRNP (Section 8.2.4), showed that addition of either 0.01% Triton X-100 or 0.05% NP40 to the buffers for the procedure was sufficient to dramatically reduce the background. NP40 can be used in splicing reactions, so it was chosen for further affinity purifications.

After suitable conditions for traditional affinity purification were determined, the photocleavage procedure was developed using GGA PC Bio. As the affinity purification procedure up to the point of UV cleavage was done in the dark to avoid unwanted cleavage, the washing steps were done using columns with frits to hold the beads. This introduced the challenge of applying UV to the samples on the column. This was overcome by assembly of a homemade contraption capable of holding the columns in a chilled environment, with a small glass plate to block out any unwanted short wave

UV, as this could cause unwanted crosslinking. The Perspex holder for the 365 nm UV wands (UVP Bio-Lite) was placed on the glass plate (Figure 3.3). Unfortunately, the spacing of the column holder and the UV wand holder were not the same, so only two samples could be cleaved simultaneously.

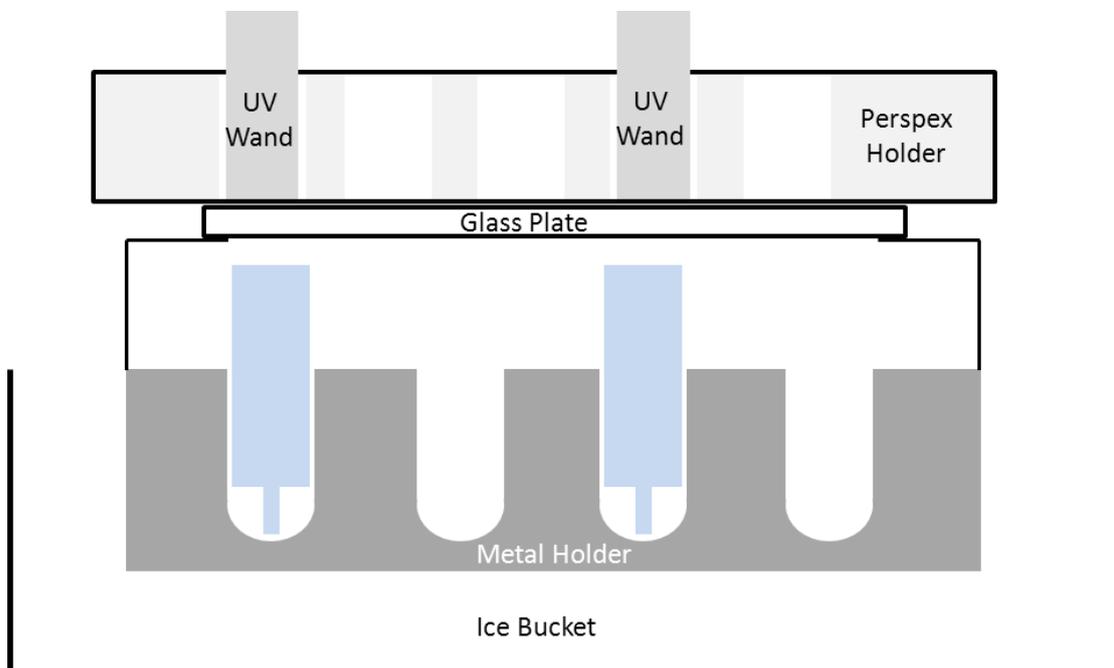


Figure 3.3: Diagram of the assembly used for UV application to affinity purified samples in columns.

The columns (blue) were used to contain the Neutravidin bead, facilitating easy and thorough washes under minimal red light conditions without losing any beads. Subsequent release of the purified samples by photocleavage using 365 nm UV was achieved on the column while the beads were chilled and moist (but not suspended in liquid), using the assembly shown above. A small glass plate was placed between the samples and the UV wands to block out any unwanted short wavelength UV that might have caused unwanted crosslinks to form.

Conditions for irradiation with 365 nm UV were optimized using 5' radiolabelled GGA PC Bio. Three samples were tested with different patterns of 365 nm UV irradiation totaling 10 minutes. The first sample had 10 minutes of UV followed by a 400 μ l elution. The second had two 5 minute applications of UV with a 200 μ l elution in-between and one at the end. The third had four 2.5 minute applications of UV with 100

μ l elutions in-between and one at the end. As shown in Figure 3.4, 10 minutes was sufficient for complete cleavage. The procedure involving four 2.5 minute UV applications was chosen for future experiments, as it allowed for more thorough elution in the same volume and kept the sample more moist throughout.

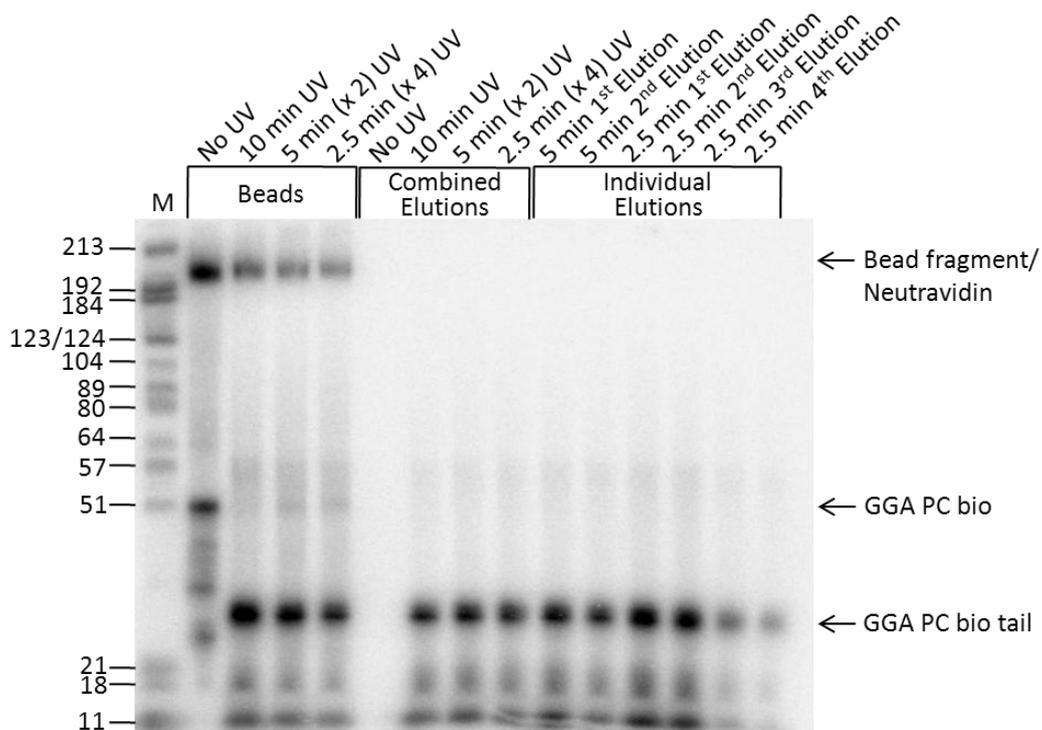


Figure 3.4: Optimization of photocleavage of 5' radiolabelled GGA PC bio while attached to Neutravidin beads, using set-up described in Figure 3.3.

This denaturing polyacrylamide gel shows that 4 consecutive 2.5 min UV applications followed by 100 μ l elutions increases the yield of eluted GGA PC Bio tail compared to one application of UV for 10 min followed by a 400 μ l elution. Also, application of 365 nm UV light for 10 min is sufficient to completely cleave the GGA PC Bio.

The optimizations of the affinity purification procedure and the UV cleavage process were combined to obtain very clean samples for mass spectrometry, the results of which are discussed in Chapter 4. As shown in Figure 3.5, the photocleavage procedure allows for selective elution of proteins associated with the enhancer tail of GGA PC bio.

This method also avoids contamination with NeutrAvidin® and non-specific proteins bound to the beads, which are released into samples that have been heated for elution (Figure 3.5).

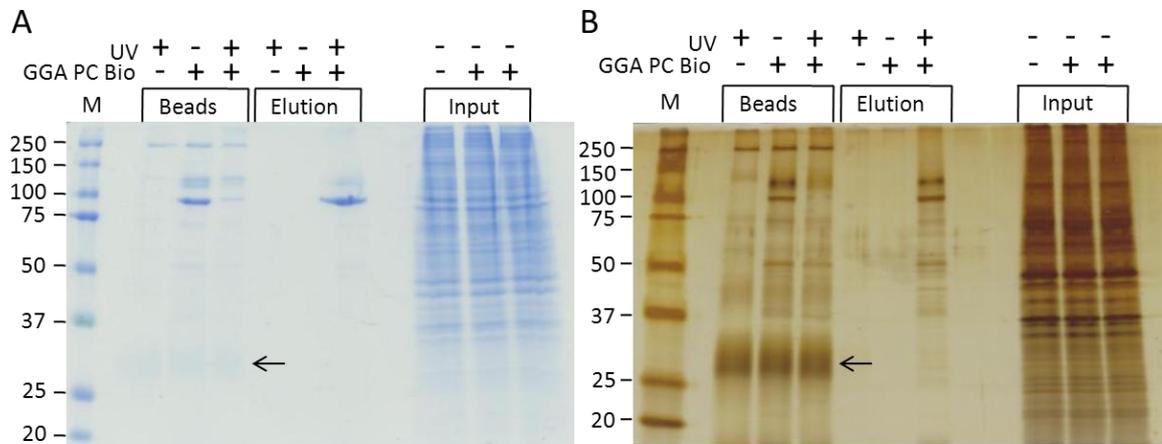


Figure 3.5: Affinity purification using a photocleavable linker improves the purity of eluted samples.

(A) Coomassie stained SDS PAGE gel of affinity purification from HeLa nuclear extract using GGA PC Bio shows that background is reduced after elution, and contamination with neutravidin (arrow) is eliminated. **(B)** Silver stained SDS PAGE gel of the same samples in Figure 3.5A shows with greater sensitivity the increased purity after elution by photocleavage. Arrow indicates the large amount of neutravidin that is left behind on the beads after elution.

3.4 Extension of Method to Purify Intact Complexes

The next goal was to use the gentle cleavage conditions provided by the photocleavage elution method for biotin affinity purification to elute an intact complex on the tail of GGA PC Bio. Native polyacrylamide gel electrophoresis of 5' radiolabelled oligonucleotides was used to assess the eluted complex, allowing for comparison with the complex that is formed before purification. As show in Figure 3.6A, a complex of the expected size was successfully purified and eluted intact using the optimized

photocleavage method. One limitation of the method was the inevitable dilution of the complex after purification. However, the use of microconcentrators allowed the complex to be concentrated significantly (Figure 3.6B).

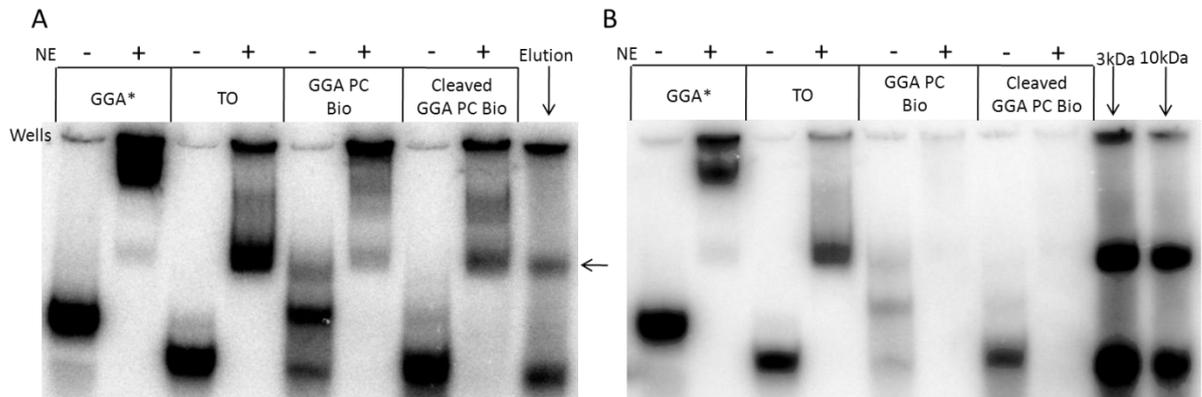


Figure 3.6: The gentle photocleavage method of elution using GGA PC Bio allows for elution of an intact complex on the GGA tail. (A) Native polyacrylamide gel using 5' radiolabelled oligonucleotides and HeLa nuclear extract shows that the affinity purified GGA PC Bio tail complex (arrow) migrates at the same size as the TO and unpurified, pre-cleaved GGA PC Bio complexes. **(B)** The eluted tail complex can be concentrated using either 3 kDa or 10 kDa MW cut-off microconcentrators (Millipore).

After successful purification of an intact small complex on the tail of GGA PC Bio, efforts were turned to purification of a much larger complex, spliceosomal complex A. Despite not being designed for this purpose, GGA PC Bio was used for financial reasons. The pre-mRNA used was β -globin exons 2-3 (Section 8.1.3), a traditional splicing substrate. A 23 nt extension on the 3' end of the β -globin construct was necessary to provide a binding site for the tail of GGA PC Bio (Section 8.1.4). *In vitro* splicing and complex formation checks of the modified β -globin construct were performed in the presence of 250 nM GGA PC Bio to ensure that the modifications were not disruptive (Figure 3.7A and C). Affinity purification with photocleavage was then attempted on samples that were stalled in spliceosomal A complex using an α -U6 snRNP oligonucleotide (Dönmez *et al.*, 2007). The procedure was successful in

selectively purifying some sort of complex, but it was not A complex. It seems from the control samples, which had to be diluted because of the weakness of the eluted sample, that A complex was not durable enough to withstand the dilutions necessary for this procedure (Figure 3.7B). Further optimizations were not attempted. However, it seems possible that this method, with some optimizations, could be useful for the purification of large complexes that are stable enough to withstand the necessary dilution.

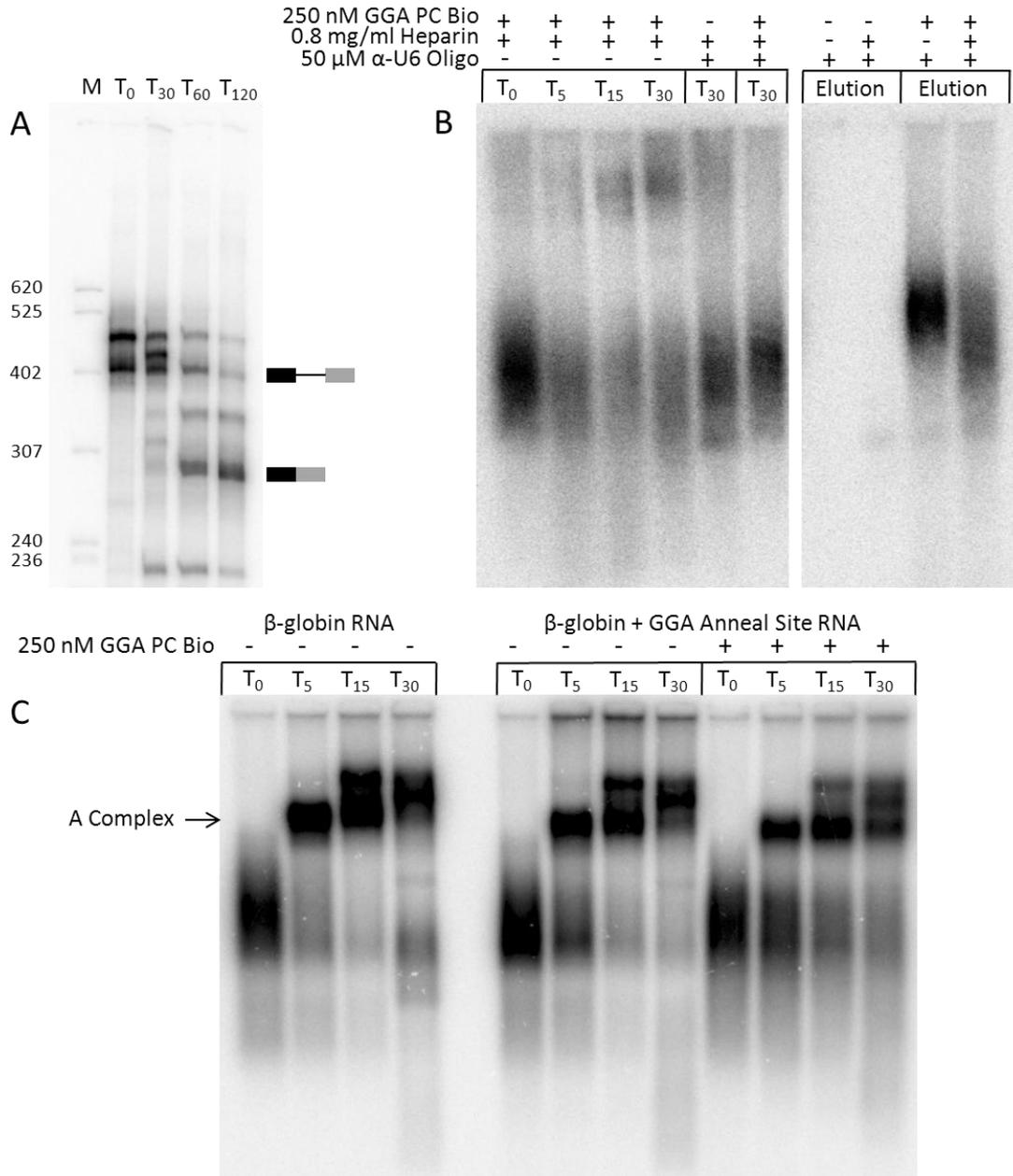


Figure 3.7: Affinity purification of β -globin splicing complex A using GGA PC Bio annealed at the 3' end of β -globin RNA was not successful (A) Denaturing PAGE of *in vitro* splicing in HeLa nuclear extract of radiolabelled β -globin + GGA anneal site (Section 8.1.4) with GGA PC Bio shows that the addition of a binding site for GGA PC Bio on the 3' end of the β -globin RNA did not inhibit splicing. (B) Native agarose gel with radiolabelled β -globin RNA (Section 8.1.3) shows the attempted affinity purification of β -globin splicing complex A using GGA PC Bio. Unfortunately, even the control samples, which had to be diluted to a level of radioactivity comparable to the eluted sample, did not remain as well-formed splicing complexes. Some complex was cleanly eluted, but it was not A complex. (C) Native agarose gel with radiolabelled RNA shows what the undiluted splicing complexes should look like and proves that the modifications made to β -globin did not inhibit formation of splicing complex A.

3.5 Discussion

Traditional biotin affinity purification is a very useful method with many applications.

This method was improved by the introduction of a photocleavable linker into the biotinylated substrate. The gentle elution conditions provided by photocleavage allowed for a stable complex on the enhancer tail of GGA PC bio to be eluted intact from NeutrAvidin® beads after purification for further analysis. This elution method also greatly reduced the contamination of samples with non-specifically bound proteins and NeutrAvidin® released during traditional harsh heating methods of elution. The increased purity of the eluted samples allowed for lower, more biologically relevant, concentrations of biotinylated substrate to be used.

Concentrations as low as 50 nM were successfully tested (data not shown), which is a great improvement in comparison with typical concentrations used, such as 1 µM. The photocleavage method of elution makes biotin affinity purification a much more powerful tool.

A range of photocleavable linkers are now commercially available, extending the applications of traditional biotin affinity purification. Variable positioning of the photocleavable linker in DNA/RNA allows for a range of new applications. Also, biotinylated photocleavable linkers can be added to primary amines of proteins, extending the benefits of photocleavage to experiments with protein substrates. Photocleavable linkers can even be used in oligonucleotide-peptide conjugates.

Traditional biotin affinity purification is a robust method with wide applications, and the improvements provided by photocleavable linkers make this method even more powerful and versatile.

4 MECHANISMS OF ACTION OF A BIFUNCTIONAL OLIGONUCLEOTIDE ENHANCER OF SPLICING

4.1 Introduction

4.2 Complexes and Bound Proteins

4.3 G-Quadruplex Nature of the GGA Tail

4.4 Actions of hnRNP F

4.5 Importance of Oligonucleotide Chemistry

4.6 Analysis of the Bifunctional Nature of GGA

4.7 Discussion

4.1 Introduction

GGA is a targeted oligonucleotide enhancer of splicing (TOES), which was developed for the treatment of spinal muscular atrophy (Skordis *et al.*, 2003). Spinal muscular atrophy (SMA) is a genetic disease resulting from the loss of the SMN1 gene (Pearn, 1980; Lunn & Wang, 2008; Feldkötter *et al.*, 2002). Fortunately, SMA patients have at least one copy of SMN2, a very similar gene that only differs by a few translationally silent mutations. However, these mutations, especially a C to T transition at position 6 in exon 7 (C6T), lead to skipping of exon 7 in 90% of transcripts, causing production of a truncated SMN protein that is rapidly degraded (Monani *et al.*, 1999; Lorson *et al.*, 1999; Le *et al.*, 2000; Cartegni *et al.*, 2006; Lorson *et al.*, 1998). This C6T mutation generates an exonic splicing silencer (ESS), which has been shown to bind hnRNP A1 and/or Sam68, and disrupts an exonic splicing enhancer known to bind SRSF1 (Lorson *et al.*, 1999; Monani *et al.*, 1999; Cartegni & Krainer, 2002; Cartegni *et al.*, 2006; Kashima & Manley, 2003; Kashima *et al.*, 2007; Pedrotti *et al.*, 2010). GGA targets the C6T mutation site by annealing from position 2-16 on SMN2 exon 7, and it was designed to recruit activator proteins with a non-annealed enhancer tail extending from position 16 (AGGAGGACGGAGGACGGAGGACA). Although this sequence has shown success both *in vitro* and *in vivo* (Skordis *et al.*, 2003; Meyer *et al.*, 2009), further investigation into the mechanism by which GGA restores SMN2 exon 7 inclusion was necessary in order to improve its therapeutic potential and provide knowledge that could be beneficial for the design of future TOES.

4.2 Complexes and Bound Proteins

The first step to understanding the mechanism of action of GGA was to determine which proteins were being recruited by its enhancer tail. The enhancer sequence of the tail was designed with the consensus binding site for SRSF1 (GAAGARR) in mind (Sanford *et al.*, 2009). However, it was necessary to prove that this protein was indeed recruited. Investigation into protein binding on GGA revealed a more complex situation than previously expected.

4.2.1 Identification of Proteins

The modified biotin affinity purification procedure, allowing the GGA tail and associated proteins to be eluted selectively with 365 nm UV light (Chapter 3), was used to gather a large enough sample for identification of the proteins by mass spectrometry. The entire sample, as well as a control sample lacking GGA PC bio, was run about 5 mm into an SDS PAGE gel. Then the whole, unseparated protein bands for the GGA PC bio sample and for the control, were submitted to the University of Leicester's Protein and Nucleic Acid Chemistry Laboratory (PNAACL) for trypsin digest and analysis by liquid chromatography - tandem mass spectrometry (LC MS/MS) using the LQT Orbitrap velos mass spectrometer (Thermo). This sensitive method identified, with 100% protein identification probability, 42 proteins unique to the +GGA PC Bio sample (Figure 4.1). The list of proteins contains many splicing regulatory proteins as well as a few core splicing components, such as U1A and SF3B2. Also abundant in the list are transcriptionally related proteins. It is clear from this list that the enhancer tail of GGA is capable of recruiting, or associating to, a spliceosomal complex. It seems likely given the large number of proteins found that, although no SMN2 RNA was

added to this affinity purification, GGA was able to associate, probably through protein-protein interactions, with transcripts present in the nuclear extract. While this is not a quantitative method, ranking the identified proteins by percent coverage may give some indication of relative abundance (Silva *et al.*, 2006). Notably, SRSF1, ranked 9th overall, was the highest ranked splicing activator protein. However, it was necessary to investigate further which proteins were bound directly to GGA. UV light of 254 nm was used to covalently crosslink bound proteins to 5' radiolabelled GGA and an oligonucleotide consisting of only the enhancer tail sequence of GGA, named TO (Section 8.2.1, Figure 4.2B). Further investigations, discussed later in this section, led to the indicated protein assignments. Figure 4.2A is a silver stained SDS PAGE gel of a small aliquot of the affinity purification sample sent for the analysis that produced Figure 4.1. Possible protein assignments have been made in Figure 4.2A, based on the work covered in this chapter. However, it is important to note that no specific bands were sent for identification by mass spectrometry from this experiment.

Rank	Identified Proteins	Accession Number	MW	Percent Coverage
1	Isoform 4 of Cellular nucleic acid-binding protein (CNBP)	P62633-4	20 kDa	60%
2	CDKN2A-interacting protein (CDKN2AIP)	Q9NXV6	61 kDa	58%
3	Heterogeneous nuclear ribonucleoprotein A1 (HNRNPA1)	F8VRQ1 (+4)	33 kDa	55%
4	Probable ATP-dependent RNA helicase DHX36 (DHX36)	Q9H2U1	115 kDa	53%
5	5'-3' exoribonuclease 2 (XRN2)	Q9H0D6	109 kDa	53%
6	Heterogeneous nuclear ribonucleoprotein H, N-terminally processed (HNRNPH1)	G8JLB6 (+1)	51 kDa	51%
7	Heterogeneous nuclear ribonucleoproteins A2/B1 (HNRNPA2B1)	P22626 (+1)	37 kDa	51%
8	Heterogeneous nuclear ribonucleoprotein H2 (HNRNPH2)	P55795	49 kDa	47%
9	Serine/arginine-rich splicing factor 1 (SRSF1)	Q07955	28 kDa	46%
10	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 (PLOD1)	Q02809	84 kDa	41%
11	Heterogeneous nuclear ribonucleoprotein U-like protein 1 (HNRNPUL1)	B7Z4B8 (+3)	86 kDa	40%
12	Heterogeneous nuclear ribonucleoprotein U (HNRNPU)	Q00839	91 kDa	40%
13	Interleukin enhancer-binding factor 3 (ILF3)	C9JFV5 (+7)	83 kDa	40%
14	Heterogeneous nuclear ribonucleoprotein A3 (HNRNPA3)	E7ERJ4 (+2)	34 kDa	37%
15	RNA-binding protein 45 (RBM45)	Q8IUH3 (+1)	54 kDa	37%
16	Nucleolin (NCL)	P19338	77 kDa	33%
17	Transcriptional activator protein Pur-alpha (PURA)	Q00577	35 kDa	32%
18	Nuclease-sensitive element-binding protein 1 (YBX1)	P67809	36 kDa	32%
19	Heterogeneous nuclear ribonucleoprotein F (HNRNPF)	P52597	46 kDa	30%
20	Interleukin enhancer-binding factor 2 (ILF2)	Q12905	43 kDa	28%
21	Transcriptional activator protein Pur-beta (PURB)	Q96QR8	33 kDa	28%
22	RNA-binding protein FUS (FUS)	B4DR70 (+2)	45 kDa	27%
23	Serine/arginine-rich splicing factor 2 (SRSF2)	Q01130 (+1)	25 kDa	22%
24	Serine/arginine-rich splicing factor 9 (SRSF9)	Q13242	26 kDa	22%
25	DNA-binding protein A (CSDA)	P16989 (+2)	40 kDa	21%
26	Serine/arginine-rich splicing factor 5 (SRSF5)	Q13243	31 kDa	20%
27	HCG2044799 (hCG_2044799)	H3BQZ7 (+1)	85 kDa	19%
28	Small nuclear ribonucleoprotein Sm D3 (SNRPD3)	B4DJP7 (+1)	13 kDa	16%
29	TAR DNA-binding protein 43 (TARDBP)	Q13148 (+1)	45 kDa	15%
30	U1 small nuclear ribonucleoprotein A (SNRPA)	P09012	31 kDa	13%
31	Serine/arginine-rich-splicing factor 7 (SRSF7)	C9JAB2 (+4)	27 kDa	13%
32	Heterogeneous nuclear ribonucleoprotein Q (SYNCRIP)	O60506 (+3)	70 kDa	12%
33	TATA-binding protein-associated factor 2N (TAF15)	Q92804 (+1)	62 kDa	11%
34	Heterogeneous nuclear ribonucleoprotein R (HNRNPR)	O43390	71 kDa	10%
35	Splicing factor 3B subunit 2 (Fragment) (SF3B2)	E9PJ04 (+4)	39 kDa	9%
36	Serine/arginine-rich splicing factor 6 (SRSF6)	Q13247 (+1)	40 kDa	7%
37	5'-3' exoribonuclease 1 (XRN1)	Q8IZH2 (+1)	194 kDa	7%
38	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15 (DHX15)	F5H6K0 (+1)	90 kDa	6%
39	Heterogeneous nuclear ribonucleoprotein D-like (HNRPDL)	O14979 (+2)	46 kDa	5%
40	Probable ATP-dependent RNA helicase DDX5 (DDX5)	E7ETL9 (+1)	62 kDa	5%
41	General transcription factor II-I (GTF2I)	B4DH52 (+4)	112 kDa	2%
42	Thyroid hormone receptor-associated protein 3 (THRAP3)	Q9Y2W1	109 kDa	2%

Figure 4.1: List of specific proteins found to be associated with the GGA enhancer tail after biotin affinity purification from HeLa nuclear extract using with GGA PC Bio and the photocleavage method of elution (See Chapter 3). Proteins were identified by LC MS/MS using a LQT Orbitrap velos mass spectrometer (Thermo) and ranked according to percentage peptide coverage (Silva *et al.*, 2006).

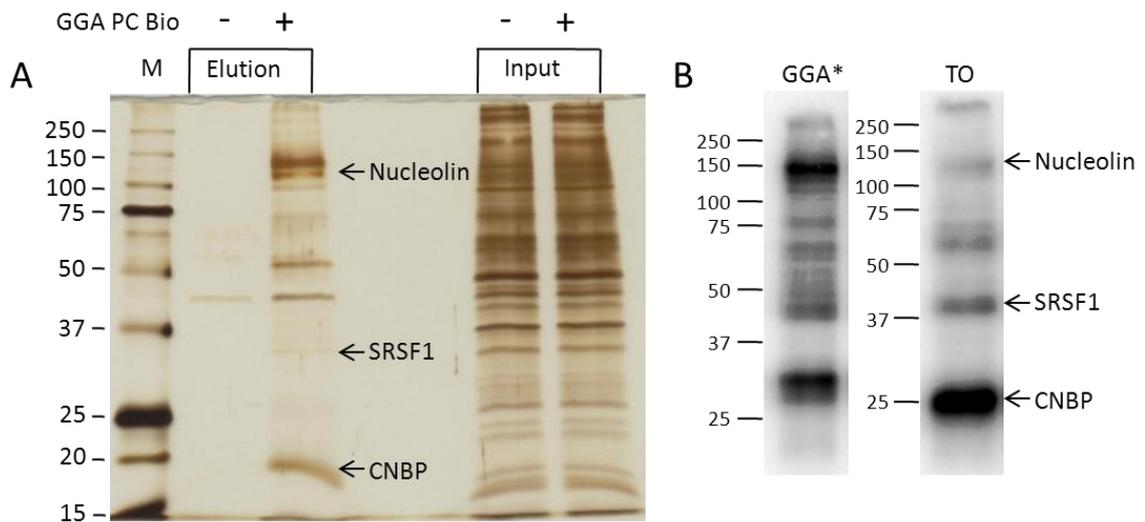


Figure 4.2: Comparison of proteins found to be associated with the tail of GGA PC Bio by affinity purification with proteins found to bind directly to GGA and the GGA tail (TO). (A) Silver stained SDS PAGE gel of a small amount of the GGA PC Bio affinity purification sample, analyzed in Figure 4.1, with potential protein assignments. (B) SDS PAGE gel transferred to nitrocellulose shows the proteins bound directly to 5' radiolabelled GGA and TO in HeLa nuclear extract using UV crosslinking. Potential protein assignments have been made.

Figure 4.3 represents a large scale biotin affinity purification where individual bands from the Coomassie stained gel (Figure 4.3A) were sent for identification by MALDI-TOF mass spectrometry after in-gel trypsin digest, performed by PNAFL. Figure 4.3B is a silver stained SDS PAGE gel of a small proportion of the same samples used in Figure 4.3A. Only two bands, indicated by arrows, were identified successfully. The top band was identified as a 5'-3' exoribonuclease, and the lower band was identified as nucleolin. It is clear from the silver stained gel that many other proteins were purified as well, but nucleolin was by far the most abundant. This result was somewhat puzzling as nucleolin is not a known splicing activator. It is important to note that the size of nucleolin did correspond well with the largest protein found by UV crosslinking (Figure 4.2B) to bind directly to GGA and TO, bearing in mind that the size of the

crosslinked oligos, 38 nt and 23 nt respectively, affects the mobility. It was not clear how either of these proteins could be responsible for the stimulatory effect of GGA on SMN2 exon 7 inclusion. Therefore, it was necessary to investigate further the other, less abundant, proteins recruited by the GGA tail.

It was unfortunate that in the experiment for Figure 4.3 the gels were run too far to detect CNBP, which we know must have been present from the Orbitrap mass spec experiment (Figure 4.1). The size of CNBP corresponds well with the smallest protein found to bind directly to GGA and TO (Figure 4.2B). To further confirm CNBP binding, a splicing functional nuclear extract was made from HEK 293T cells expressing T7-tagged CNBP (Figure 4.4A and B). Figure 4.4C compares UV crosslinking of GGA and TO in the control 293T extract to crosslinking in the T7-CNBP 293T NE. Unfortunately, the T7 tag was quite small (MASMTGGQMG), so it was difficult to detect the shift in size. For GGA it was impossible to detect the shift produced by the T7 tag because of the additional size of the crosslinked oligo. However, a small shift can be observed for TO (Figure 4.4C). Additionally, for both GGA and TO the relevant crosslinked band is brighter in the overexpressed T7-CNBP NE than the control, also implicating CNBP as the bound protein.

The binding of nucleolin and CNBP to the tail of GGA was a strong indication that the tail could be forming a G-quadruplex, as both nucleolin and CNBP are known to bind sequences which can form G-quadruplexes (Armas *et al.*, 2008; Michelotti *et al.*, 1995; Liu *et al.*, 1998; Borgognone *et al.*, 2010; Bates *et al.*, 1999; González *et al.*, 2009). NMR and CD studies on the GGA tail sequence done by colleagues from the University of

Nottingham (Mr A. Cousins, Prof. M. Searle), with the help of Mr A. Perrett (University of Leicester), proved that the GGA tail does have the ability to form a G-quadruplex (Smith *et al.*, manuscript submitted). As shown in Figure 4.2B, nucleolin binding is more prevalent with GGA than TO. This indicates that nucleolin may bind at the junction between the G-quadruplex and the annealing region, which is not part of the G-quadruplex. This type of binding has been observed previously for another G-quadruplex binding protein, FMRP (Phan *et al.*, 2011; Ramos *et al.*, 2003). Further investigation into the G-quadruplex nature of the GGA tail can be found in Section 4.3.

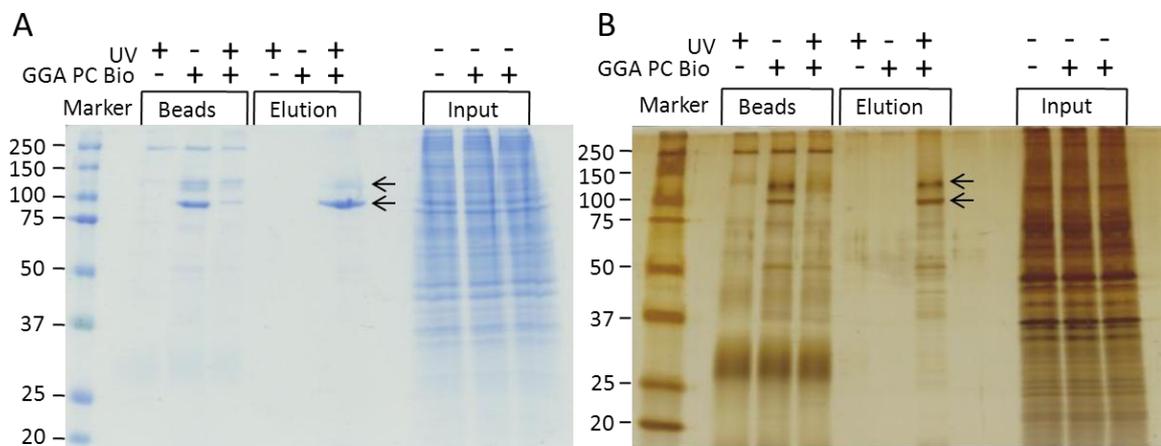


Figure 4.3: Affinity purification with GGA PC Bio revealed 5'-3' exonuclease 2 (top arrow) and nucleolin (bottom arrow) binding. (A) Coomassie stained SDS PAGE gel of affinity purification with GGA PC bio from HeLa nuclear extract. Individual bands were sent for MALDI-TOF mass spectrometry, two of which were identified successfully. **(B)** Silver stained SDS PAGE gel of the same samples from Figure 4.3A.

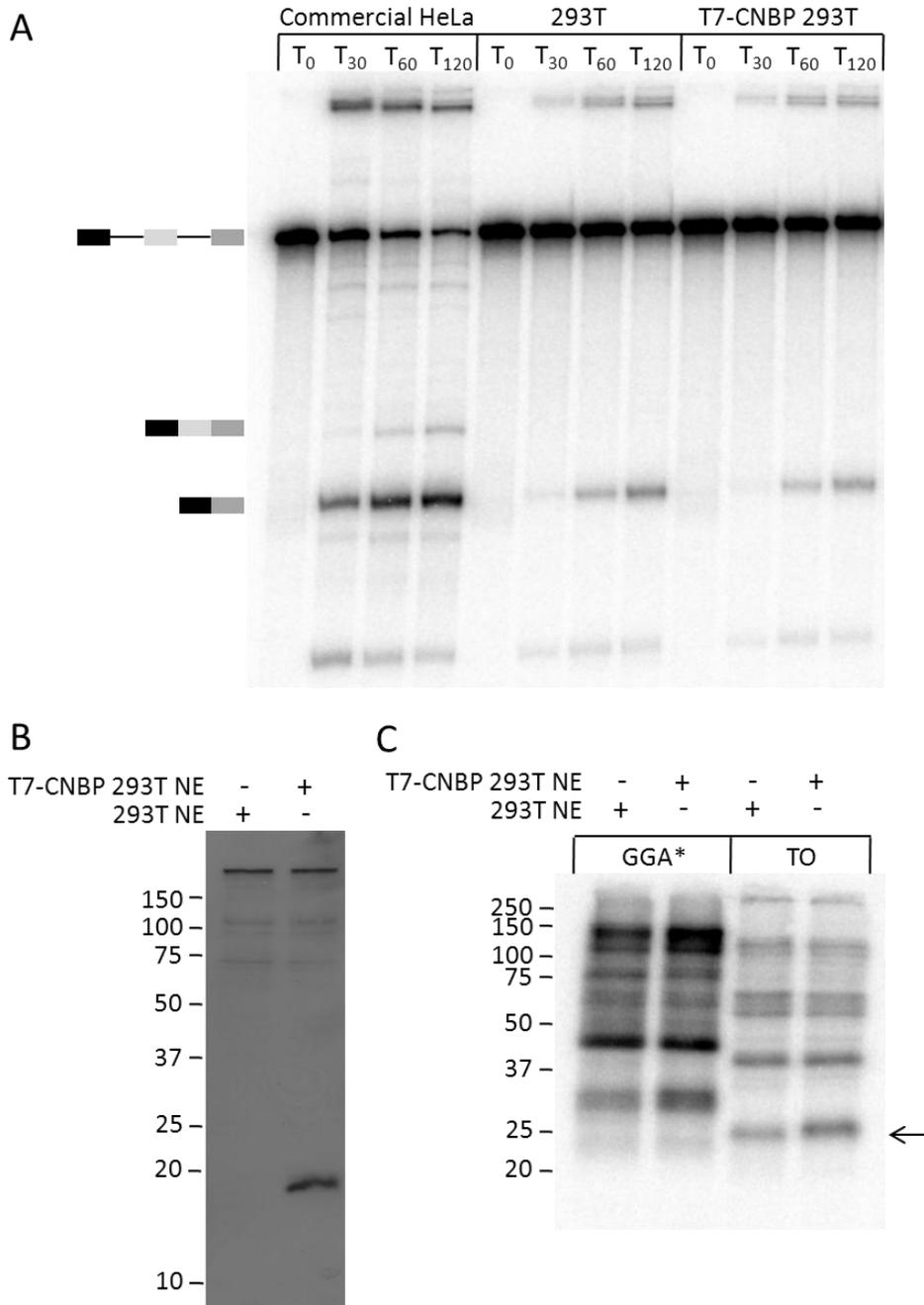


Figure 4.4: CNBP binds directly to the enhancer tail of GGA. **(A)** Denaturing PAGE of *in vitro* splicing with radiolabelled SMN2 RNA showing that the T7-CNBP HEK 293T nuclear extract is functional. **(B)** α -T7 western blot showing expression of T7-CNBP. **(C)** SDS PAGE gel transferred to nitrocellulose, indicating T7-CNBP binding to 5' radiolabelled GGA and related tail-only oligo (TO – Section 8.2.1) by UV crosslinking. For GGA, only an increase in intensity of the CNBP band, due to binding of CNBP and the over-expressed T7-CNBP, can be observed due to the relatively small additional mass of the T7 tag. However, a shift from the T7 tag, as well as an increase in intensity, is observed when the TO oligo is crosslinked (arrow), as the additional mass of the crosslinked TO is less than for GGA.

While identification of nucleolin and CNBP as two of the main proteins bound directly to GGA may be explained by the ability of the GGA tail to form a G-quadruplex, this still did not explain the positive effect of GGA on SMN2 exon 7 splicing. As mentioned previously, GGA was designed to bind the splicing activator protein SRSF1. Also, SRSF1 was found to be associated with the tail of GGA through affinity purification (Figure 4.1). However, it was necessary to prove that SRSF1 was binding directly to GGA. This was done using UV crosslinking to 5' radiolabelled GGA and TO in nuclear extracts containing eGFP-labelled SRSF1 (Figure 4.5). Figure 4.5A-D were all derived from the same experiment where GFP-SRSF1 NE from HeLa cells was used. It is clear from the appearance of a new crosslinked band for both GGA and TO in the samples containing GFP-SRSF1 that GFP-SRSF1 is capable of binding directly to the tail of GGA (top arrows in Figure 4.5A). The bottom arrows in Figure 4.5A indicate the bands potentially corresponding to unlabelled, crosslinked SRSF1, based on size comparisons made with the anti-SRSF1 western blot (Figure 4.5C). Figure 4.5B shows the location of the bands from western blot (Figure 4.5C) when overlaid with UV crosslinking image. The observed shift in size is due to the additional size of the crosslinked oligos. It is strange that the ratio of labelled to unlabelled SRSF1 in the western blot does not seem to correspond well with the ratio of the GFP-SRSF1 crosslinked band to the proposed SRSF1 crosslinked band. However, it is clear that GFP-SRSF1 does bind. Figure 4.5D is a native polyacrylamide gel of the same samples used for Figure 4.5A-C, indicating that the relevant complexes were forming on GGA and TO in the GFP-SRSF1 NE. Figure 4.5E shows another UV crosslinking experiment where NE from HEK 293T cells expressing GFP-SRSF1 was used. In this case a more dramatic shift to binding almost exclusively GFP-SRSF1 is observed for both GGA and TO. This is likely due to the fact that

endogenous SRSF1 expression, and overexpression of GFP-SRSF1, is higher in 293T cells than HeLa cells. It also indicates that there is likely competitive binding between the quadruplex binding proteins (CNBP, nucleolin) and SRSF1. It seems likely from these findings that the stimulatory effect of GGA on SMN2 exon 7 splicing is due to direct binding of SRSF1 to the enhancer tail of GGA, followed by protein-protein interactions, ultimately resulting in recruitment of the splicing machinery.

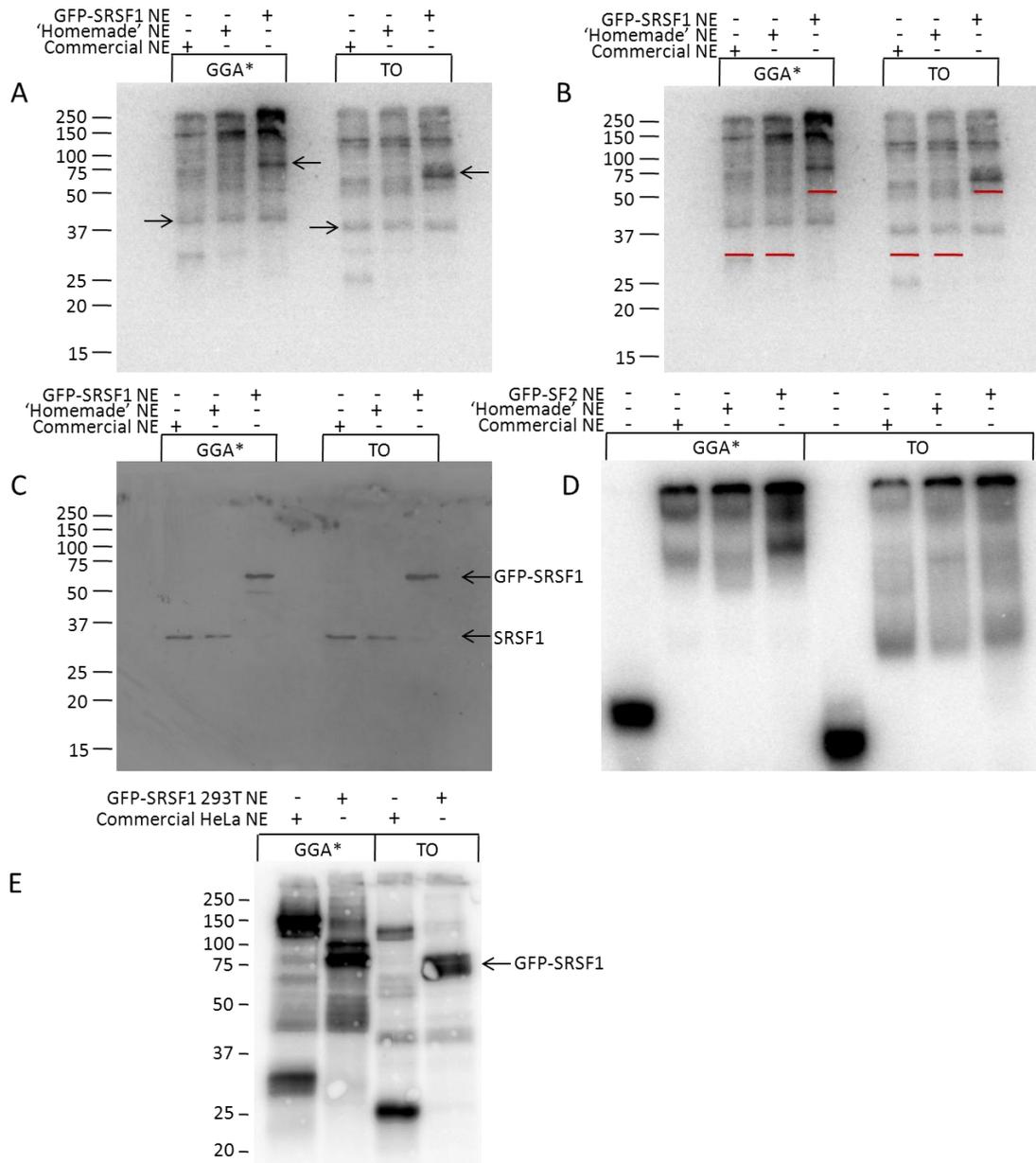


Figure 4.5: SRSF1 binds directly to the enhancer tail of GGA (TO). (A) UV crosslinking samples with 5' radiolabelled GGA and TO in HeLa nuclear extracts, including GFP-SRSF1 NE, which were run on an SDS PAGE gel transferred to nitrocellulose. The new band that appeared in the GFP-SRSF1 nuclear extract, indicated by the leftward pointing arrows, is consistent with the size expected for GFP-SRSF1. Rightward facing arrows indicate a band of the right size to be SRSF1. (B) Same image as Figure 4.5A with red lines indicating the position of the SRSF1 bands identified by western blot (Figure 4.5C). The shift in size due to the mass of the crosslinked oligos is consistent with expectations. (C) α -SRSF1 western blot of the nitrocellulose from Figure 4.5A, showing GFP-SRSF1 expression. (D) Native polyacrylamide gel of the same samples used in Figure 4.5A. (E) GFP-SRSF1 becomes the predominantly bound protein, as assessed by UV crosslinking to 5' radiolabelled GGA and TO, when highly over-expressed in HEK 293T nuclear extract.

4.2.2 Binding Specificity and Strength

As multiple proteins were found to bind to GGA, investigations into the relative binding specificity and strength of these proteins were done. Figure 4.6 shows that a 50 fold excess of GGA can successfully compete off proteins, such as nucleolin and CNBP, bound specifically to the tail of GGA PC Bio in an affinity purification experiment. Interestingly, a 50 fold excess of a different oligonucleotide, named AAC (Section 8.2.4), was unable to disrupt CNBP binding, but did seem to disrupt nucleolin binding (Figure 4.6). Unfortunately, other protein bands were too weak to analyze in this experiment.

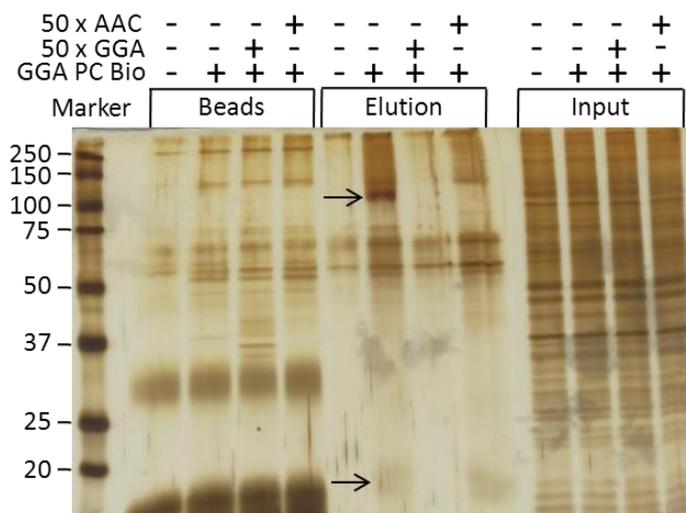


Figure 4.6: Silver stained SDS PAGE gel of affinity purification with GGA PC Bio from HeLa nuclear extract shows that addition of 50 x GGA successfully competes off the proteins specifically bound to GGA PC Bio. Competition with 50 x AAC oligonucleotide disrupted nucleolin binding (top arrow), but CNBP binding was unaffected (bottom arrow). See section 8.2 for oligonucleotide details.

Comparison of protein binding to the enhancer tail, determined by UV crosslinking, with complex formation, determined by native polyacrylamide gel electrophoresis, often reveals a similarity in pattern (Figure 4.7A and B). With all of the experiments

done to date, there is no way to be sure of the protein composition in the native gel bands. However, it seems likely that the predominant GGA tail complex band (smallest band) is due to CNBP binding. Figure 4.2B showed that CNBP was the most abundant protein crosslinked to the tail of GGA (TO). Likewise, the 2'OMe version of TO, O-TO, also shows CNBP crosslinking (Figure 4.7A). Comparison of the crosslinking pattern of proteins to O-TO (Figure 4.7A) to the complexes seen in the native gel (Figure 4.7B) indicates that the smallest complex band could be CNBP. It is therefore likely that the smallest complex band on GGA-O, the 2'OMe version of GGA, is also due to CNBP binding. Interestingly, as the concentration of nuclear extract is reduced, binding to the tail of GGA-O is favored (Figure 4.7C). It is important to note that this strong tail band, which is likely CNBP, is also found on the tail of GGA (Figure 3.6).

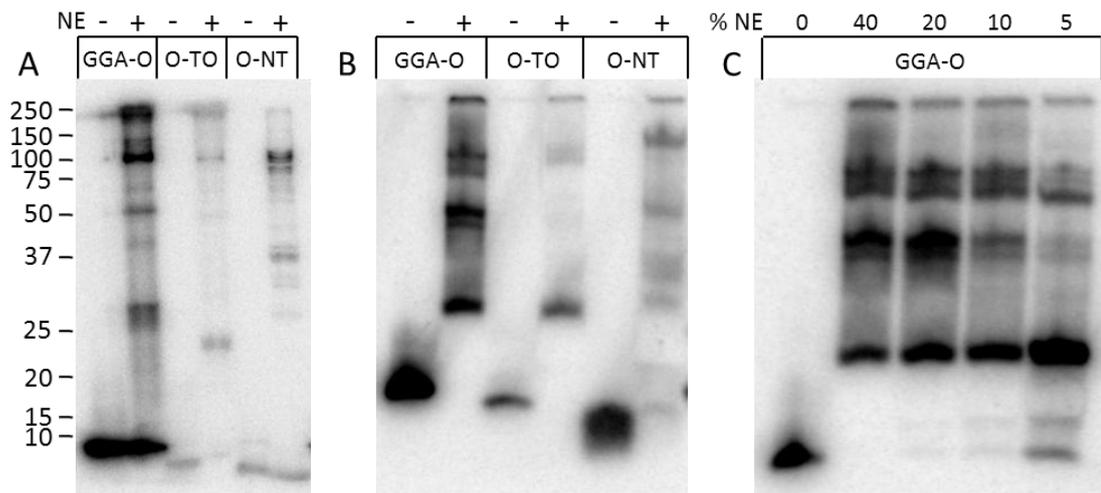


Figure 4.7: Analysis of protein binding to GGA-O, O-TO and O-NT by UV crosslinking and native polyacrylamide gel electrophoresis

(A) SDS PAGE gel of 5' radiolabelled GGA-O, O-TO and O-NT UV crosslinked in HeLa nuclear extract. (See Section 8.2.1 for oligonucleotide details.) **(B)** Native polyacrylamide gel of 5' radiolabelled oligos in HeLa nuclear extract shows a similarity to the pattern observed by UV crosslinking (Figure 4.7A), especially for O-TO. **(C)** Native polyacrylamide gel showing that as the concentration of nuclear extract is reduced, binding to the enhancer tail of GGA-O is favored, as determined by comparison with the nature of the O-TO complex in Figure 4.7B.

Native gel experiments investigating SRSF1 binding, done using T7 tagged SRSF1 purified from E.Coli, found that while SRSF1 does bind to GGA-O, binding is very sensitive to handling and electrophoresis conditions (Figure 4.8). Of the many conditions tested, only the sample incubated at 4 °C in normal splicing buffers, which was never placed at 30 °C or -80 °C, showed binding of the purified SRSF1, and even then only under one of the four tested electrophoresis conditions (Figure 4.8C). As 5% glycerol in the native gel was found to be beneficial in retaining SRSF1 binding (Figure 4.8C), 10% glycerol was subsequently tested and found to be even more beneficial, stabilizing SRSF1 binding even in samples that had been placed at room temperature or frozen at -80 °C (Figure 4.9). It is possible that the weak binding was due to the instabilities of having a purified protein. Also, any G-quadruplex structure of the tail of GGA-O is likely to have made it more difficult for the purified SRSF1 to bind.

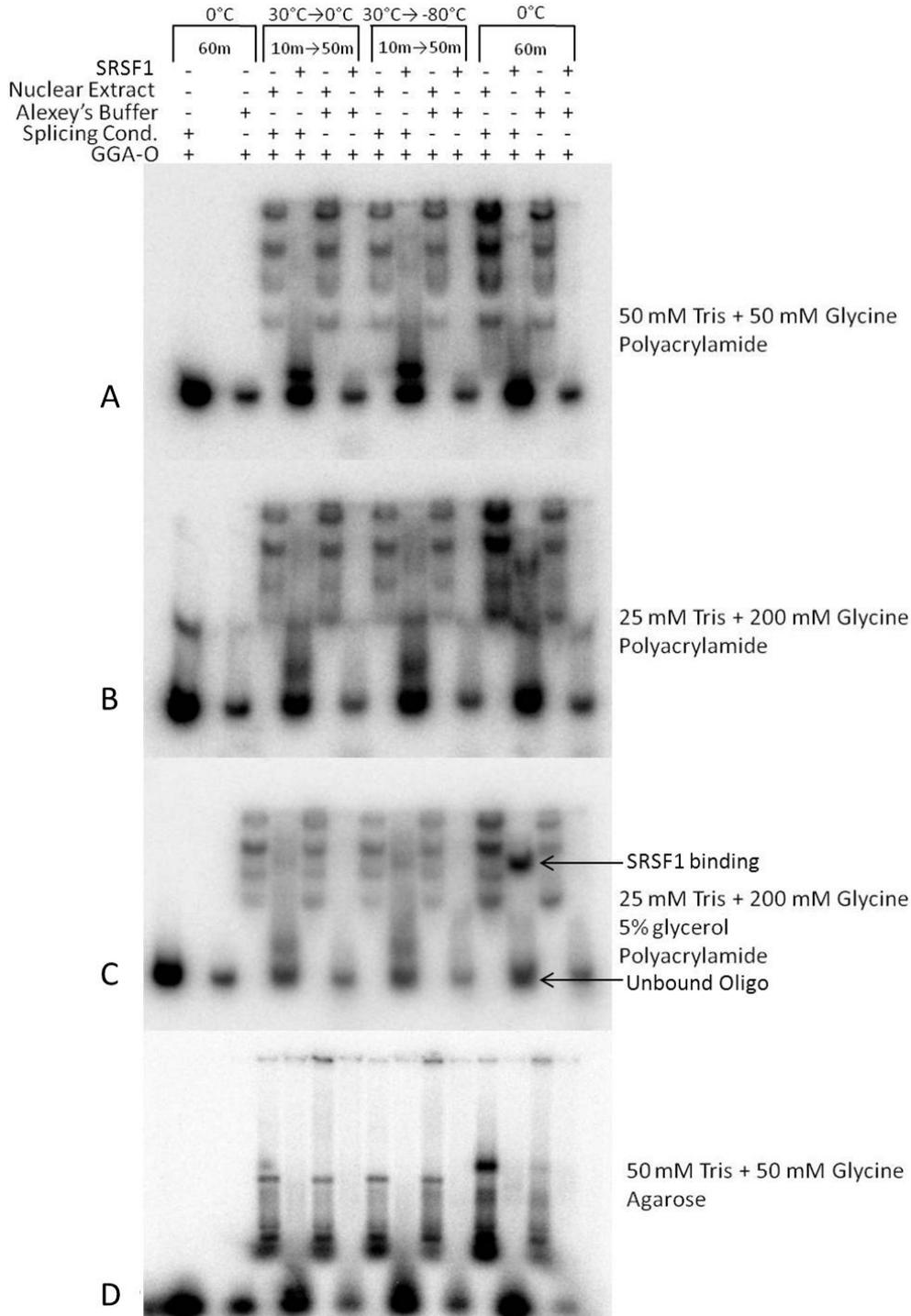


Figure 4.8: Purified SRSF1 binding to 5' radiolabelled GGA-O analyzed by native gel Binding is very sensitive to incubation conditions and electrophoresis conditions. All parts of this figure are different electrophoresis conditions for the same samples. **(A)** No SRSF1 binding retained **(B)** SRSF1 binding detected, but only the sample that was never incubated at 30 °C or frozen at -80 °C. **(C)** Addition of 5% glycerol to the native polyacrylamide gel stabilizes the SRSF1 binding observed in part B. **(D)** Native agarose gel with same buffer as part A was similarly unsuccessful in retaining SRSF1 binding.

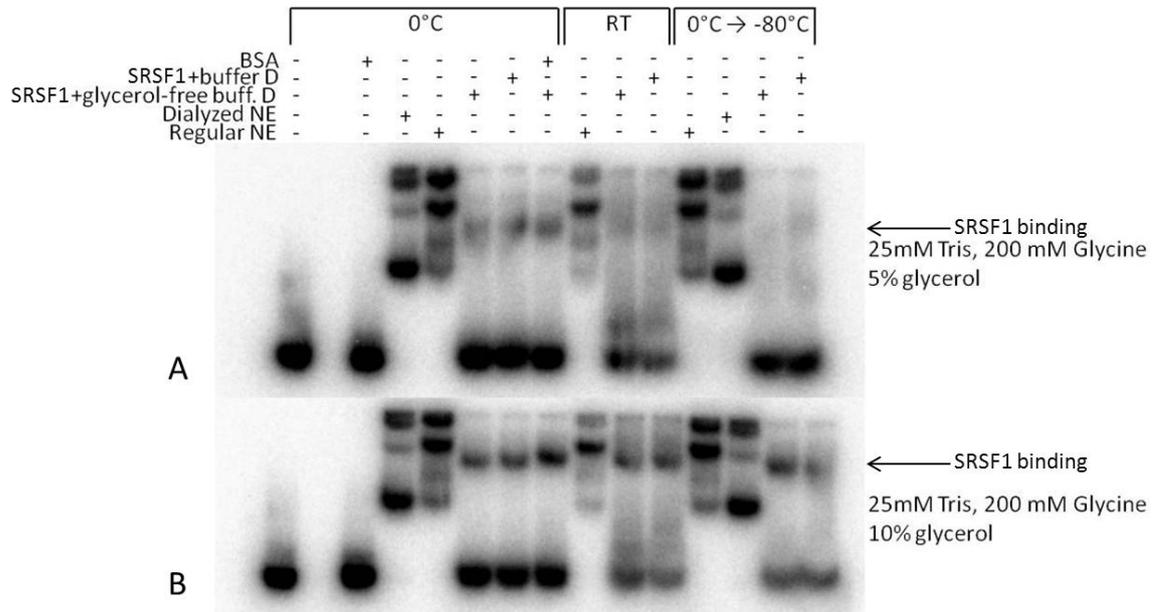


Figure 4.9: Detection of purified SRSF1 binding to 5' radiolabelled GGA-O by native polyacrylamide gel electrophoresis is further improved by increasing the amount of glycerol in the gel from 5% (A) to 10% (B).

4.2.3 Protein Binding to GGA is Unchanged in Presence of SMN2 RNA

Many experiments were done looking at protein binding to GGA under splicing conditions but without SMN2 RNA. It was necessary to determine if GGA bound the same proteins in the same ratios when in the presence of SMN2 under splicing conditions. To this end, 3' radiolabelled GGA was incubated with an approximate 2 fold excess of cold SMN2 RNA under splicing conditions, and protein binding was monitored by UV crosslinking at various timepoints (Figure 4.10B). Quantification of the ratios of the three clear crosslinked protein bands revealed a loss over time of nucleolin binding with an increase in SRSF1 binding. CNBP binding was less affected. However, the same trend was seen in the absence of SMN2 RNA (Figure 4.10A).

Therefore, this effect was attributed to some other changes, such as phosphorylation or other protein modifications, which would vary over time.

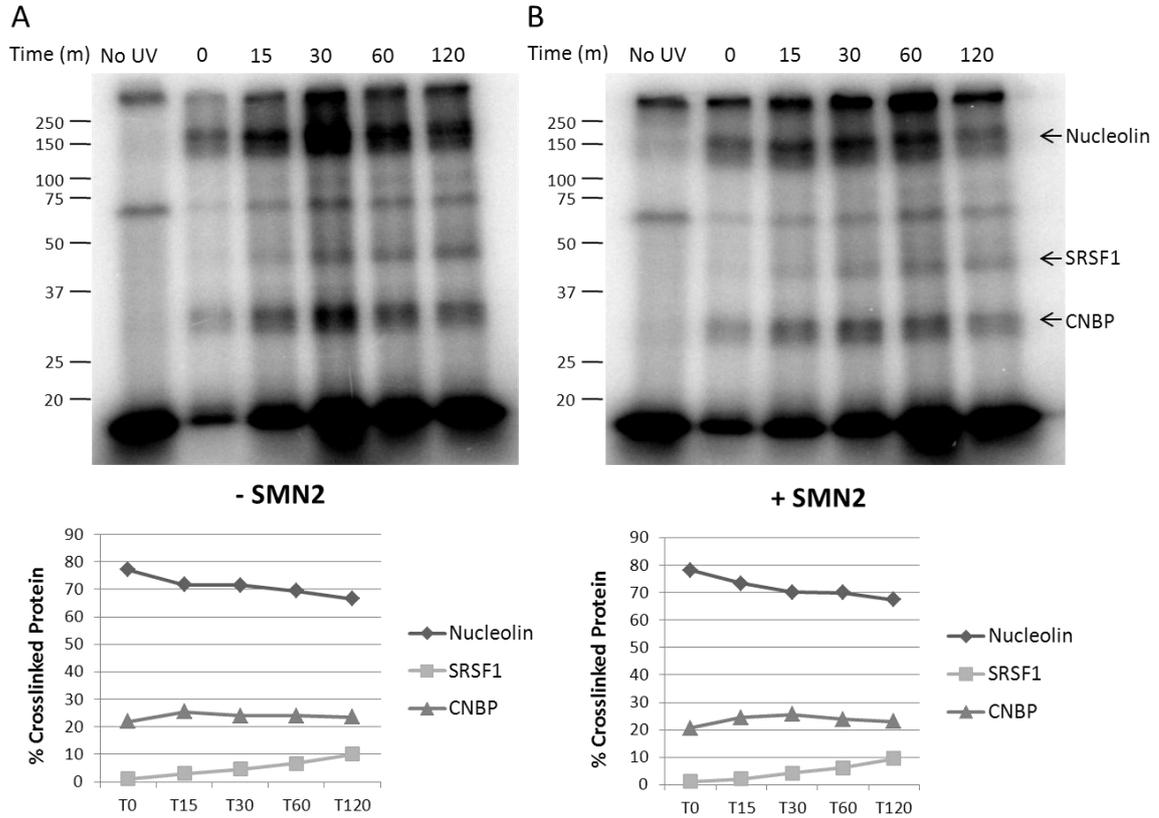


Figure 4.10: Timecourse of UV crosslinking of GGA in HeLa nuclear extract +/- SMN2

The composition of proteins bound to GGA in HeLa nuclear extract is unchanged in the presence of SMN2 RNA. However, over time in nuclear extract at 30 °C the relative level of nucleolin binding to GGA is reduced while the level of SRSF1 is slightly increased. **(A)** SDS PAGE gel of UV crosslinking over time with 3' radiolabelled GGA under splicing conditions, but without the SMN2 RNA transcript. Below is a graph of the relative percentage of the three main crosslinked proteins, determined by quantification with OptiQuant software of the intensity of radiation from each band, which was detected using a phosphorimaging screen and Cyclone reader (Section 2.2.7). **(B)** SDS PAGE gel of UV crosslinking with 3' radiolabelled GGA over time under splicing conditions with the SMN2 RNA transcript. Below is a graph of the relative percentage of the three main crosslinked proteins, determined by quantification of the intensity of radiation from each band.

In order to determine the effect of phosphorylation status on the proteins binding to

GGA, affinity purification using GGA PC Bio was done comparing normal (+ATP) and -

ATP conditions (Figure 4.11). Strikingly, nucleolin binding was greatly reduced when ATP was depleted, indicating that phosphorylation of nucleolin may be necessary for binding to GGA. For the most part, the binding of other proteins, including CNBP, was unaffected by ATP depletion. It seems likely that the reduction in nucleolin binding over time, observed in Figure 4.10, is due to a reduction in ATP levels. The increase in SRSF1 binding could be somewhat opportunistic, as nucleolin becomes unable to bind.

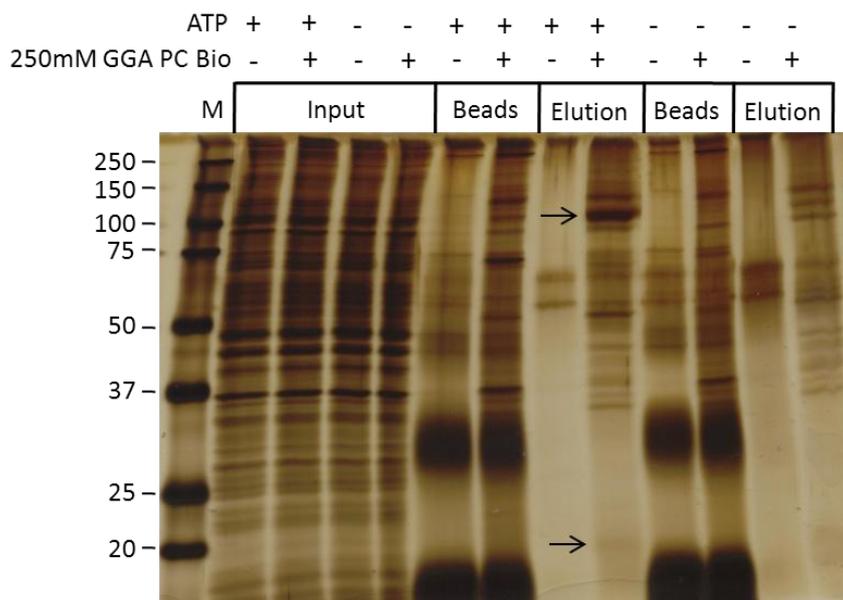


Figure 4.11: Nucleolin binding is greatly reduced when ATP is depleted.

This silver stained SDS PAGE gel of affinity purification with GGA PC Bio from HeLa nuclear extract shows that CNBP (bottom arrow) and most other protein binding is unaffected by ATP depletion conditions, but nucleolin binding (top arrow) is greatly reduced.

In an attempt to separate GGA molecules in SMN2 splicing complexes from free GGA, and to then compare the proteins bound in each state, glycerol gradients were performed on UV crosslinked samples containing 3' radiolabelled GGA with and without roughly 2 fold excess cold SMN2 RNA. In order to create a uniform sample, an oligonucleotide against the U6 snRNP was used to stall at the splicing complex A

before the crosslinking step (Dönmez *et al.*, 2007). The fractions from the gradients were measured using Cherenkov scintillation counting, before being TCA precipitated and analyzed by SDS PAGE (Figure 4.12A-C). A control gradient using radiolabelled β -globin RNA, stalled at A complex, was done simultaneously to give an indication of the expected fractionation pattern for A complex (Figure 4.12D). The fractionation and crosslinking patterns for the GGA gradients were very similar, indicating that protein binding was not significantly altered in the presence of SMN2. Unfortunately, comparison of the scintillation readings for the GGA gradients with the β -globin gradient revealed that even when SMN2 was present, GGA was not remaining annealed during the gradient process (Figure 4.12C and D). As the SMN2 construct is essentially the same as the 2 exon β -globin construct with SMN2 exon 7 and flanking sections of intron inserted, the various possible A complexes for the SMN2 construct would likely be as large as, or larger than, the β -globin A complex, which peaks in fraction 8 of the glycerol gradient. Both GGA gradients peak in the very early fractions (2-3) and have tailed off by fraction 8. Therefore, the majority of the GGA in the + SMN2 sample was not travelling through the gradient with the SMN2 complexes.

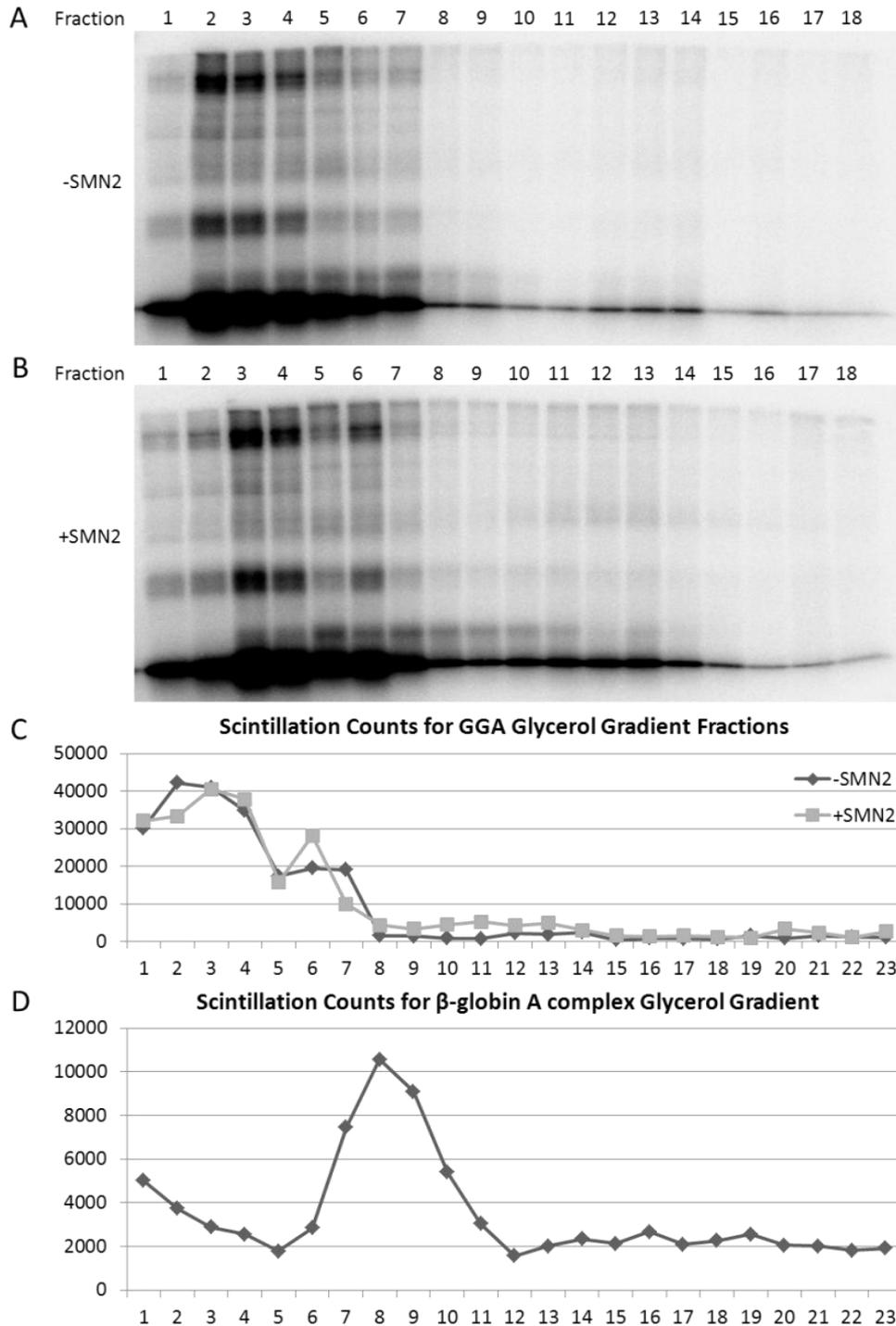


Figure 4.12: Protein binding to GGA in HeLa nuclear extract is unchanged in the presence of SMN2 RNA, which has been stalled at splicing complex A. GGA does not remain associated with the SMN2 RNA when analyzed by glycerol gradient. **(A)** SDS PAGE gel of glycerol gradient fractions from – SMN2 experiment with 3' radiolabelled GGA. **(B)** SDS PAGE gel of glycerol gradient fractions from + SMN2 experiment with 3' radiolabelled GGA. **(C)** Graph of the scintillation counts of the gradients from A and B. **(D)** Scintillation counts of a glycerol gradient with radiolabelled β -globin RNA stalled at splicing complex A.

The weak annealing of GGA observed in the glycerol gradient experiments was not entirely surprising, as the sequence is very A/U rich (Section 8.2.1). Also, many difficulties had been experienced previously during the optimization of the primer annealing step for the GGA tail reverse transcription experiments. This case is slightly different, as the GGA is annealing to a short complementary DNA sequence, but still very relevant. The GGA DNA primer would not even stay annealed to GGA-O during native polyacrylamide gel electrophoresis at room temperature. However, annealing could be detected by electrophoresis at 4 °C (Figure 4.13). GGA-L/O, a version of GGA with LNA modifications in the annealing region, was capable of stronger annealing, as would be expected. However, GGA-L/O does not enhance SMN2 exon 7 splicing as well as GGA (Section 8.2.1, (Owen *et al.*, 2011)). In fact, easy dissociation may be key to the function of GGA, as it anneals only one nucleotide away from the 5' end of SMN2 exon 7. Although the ability of GGA to readily dissociate from SMN2 may be useful for splicing, it makes many types of experiments that involve dilution unfeasible.

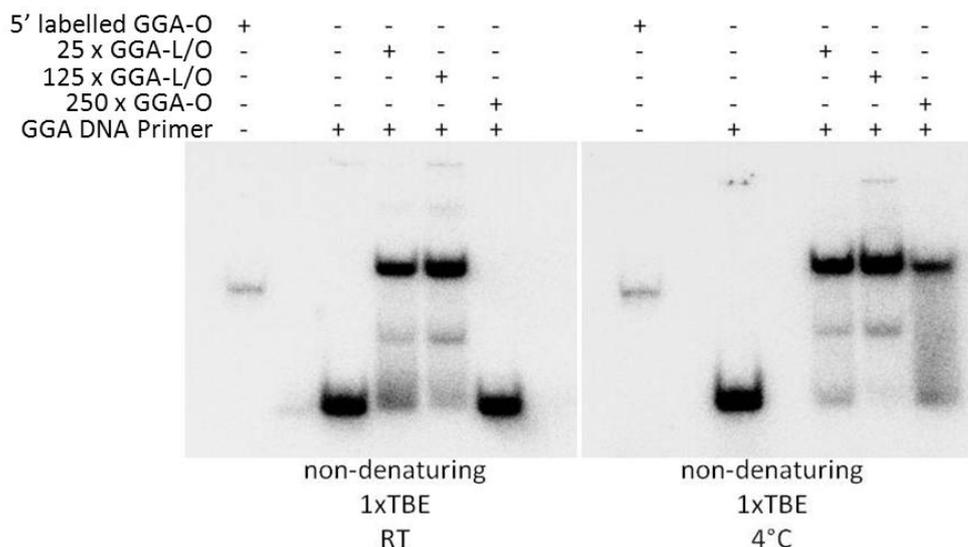


Figure 4.13: The 15 nucleotide annealing region of GGA is very weak.

A 5' radiolabelled DNA oligonucleotide complementary to the annealing region of GGA (GGA DNA primer, Section 8.2.3) would not remain annealed to GGA-O during native gel electrophoresis unless carried out at 4 °C. A version of GGA with a stronger locked nucleic acid (LNA) annealing region, GGA-L/O (Section 8.2.1) was capable of remaining annealed at room temperature.

4.3 G-quadruplex Nature of the GGA Tail

As mentioned in Section 4.2.1, the 23 nucleotide tail sequence of GGA, show in Figure 4.14, is capable of forming a G-quadruplex, based on CD and NMR studies using only the tail sequence (Smith *et al.*, manuscript submitted). This data suggested a parallel G-quadruplex, possibly involving two strands of GGA (Mr A. Cousins and Prof. M. Searle, University of Nottingham). Based on similarities of the CD data and oligonucleotide sequence to previous findings in the literature, the bimolecular G-quadruplex would likely consist of two single G-quadruplexes stacked face to face, with the intervening adenosine bases creating stabilizing hexad arrangements in conjunction with the G-tetrads (Liu *et al.*, 2002; Uesugi *et al.*, 2003; Mashima *et al.*, 2013).



Figure 4.14: Diagram of GGA oligonucleotide with G's capable of forming a G-quadruplex indicated in red.

Regardless of the exact folding structure of the G-quadruplex formed, it was necessary to investigate whether or not the G-quadruplex form of GGA was present in nuclear extract. Also, it was important to investigate whether the G-quadruplex form of GGA was the 'active' form, i.e. whether the splicing activator protein(s) responsible for the stimulation of SMN2 exon 7 inclusion binds to the G-quadruplex form of GGA or to the melted form. To this end, a series of experiments were done using three potential G-quadruplex stabilizing ligands, which were kindly provided by Prof. Laurence Hurley (University of Arizona). The structures of the three ligands GQC-05, GSA-0820 and GSA-0902 can be seen in Figure 4.15. CD analysis of the GGA tail sequence in the presence of GSA-0820 and GSA-0902 indicated that they did have a stabilizing effect (Mr A. Cousins and Prof. M. Searle, University of Nottingham, personal communication). GQC-05 was not analyzed in this way due to a lack of material.

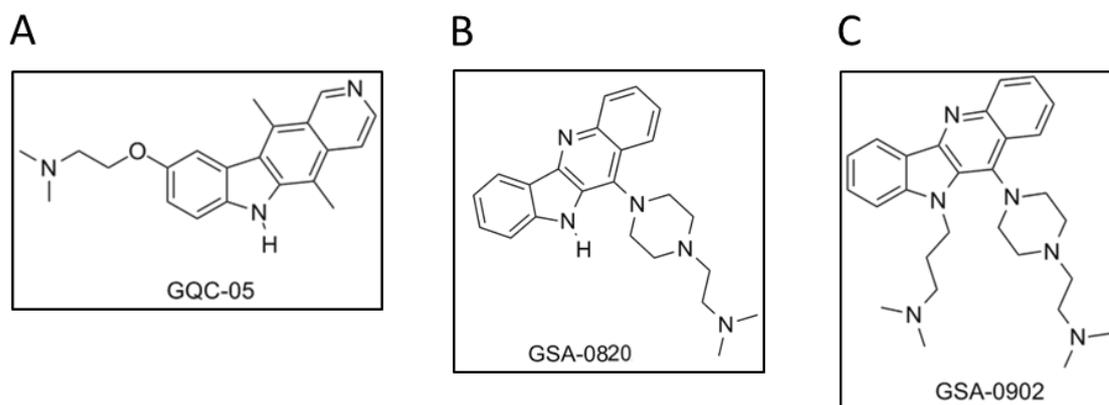


Figure 4.15: Structures of three G-quadruplex stabilizing ligands. **(A)** GQC-05 **(B)** GSA-0820 **(C)** GSA-0902

In order to investigate the role of the G-quadruplex form of GGA, a series of *in vitro* splicing assays were performed using the three G-quadruplex stabilizing ligands. Some optimization was required initially as the DMSO used to dissolve the ligands, producing a concentration of 5% DMSO in the final splicing reactions, caused a significant reduction in splicing efficiency. This was overcome by reducing the level of DMSO in the stocks to 10%, giving a final concentration of 0.5% DMSO in the splicing reactions, which did not have a significant negative effect on splicing efficiency (Figure 4.16). Also, the GQC-05 ligand caused RNA to stick to the sample tubes, especially in samples lacking the GGA oligonucleotide (Figure 4.16). This was overcome by the use of silanized plasticware for all future experiments involving the G-quadruplex stabilizing ligands. *In vitro* splicing of SMN2 with and without 250 nM GGA was done in triplicate for a series of concentrations of GQC-05: 10 times less than GGA, equimolar to GGA and 10 times greater than GGA. This showed that as the concentration of GQC-05 was increased, the level of SMN2 exon 7 inclusion decreased (Figure 4.17). The same was done for GSA-0820 and GSA-0902, and likewise SMN2 exon 7 inclusion was reduced

with increasing ligand concentration (Figure 4.18 and Figure 4.19). In all three experiments, exon 7 inclusion was also reduced somewhat in the control samples lacking GGA. In order to determine whether or not the ligands were having a more significant effect in the presence of GGA, the ratio of SMN2 exon 7 inclusion/exclusion from the samples containing ligand were normalized to the control samples lacking ligand (Figure 4.20). This showed clearly that for GSA-0820 and GSA-0902 the ligands were having a more significant effect in the presence of GGA (Figure 4.20B and C). However, for GQC-05, the effect was not as clear, and if anything seemed to be more significant in the samples lacking GGA. The difference in behaviour of GQC-05 can be explained in part by its difference in structure. GSA-0820 and GSA-0902 are very similar in structure to each other. In conclusion, the *in vitro* splicing experiments with GSA-0820 and GSA-0902 suggest that stabilization of the G-quadruplex structure of the tail of the GGA oligonucleotide has a negative effect on its ability to stimulate SMN2 exon 7 inclusion, indicating that the relevant activator protein(s) probably bind to the melted form of the GGA tail.

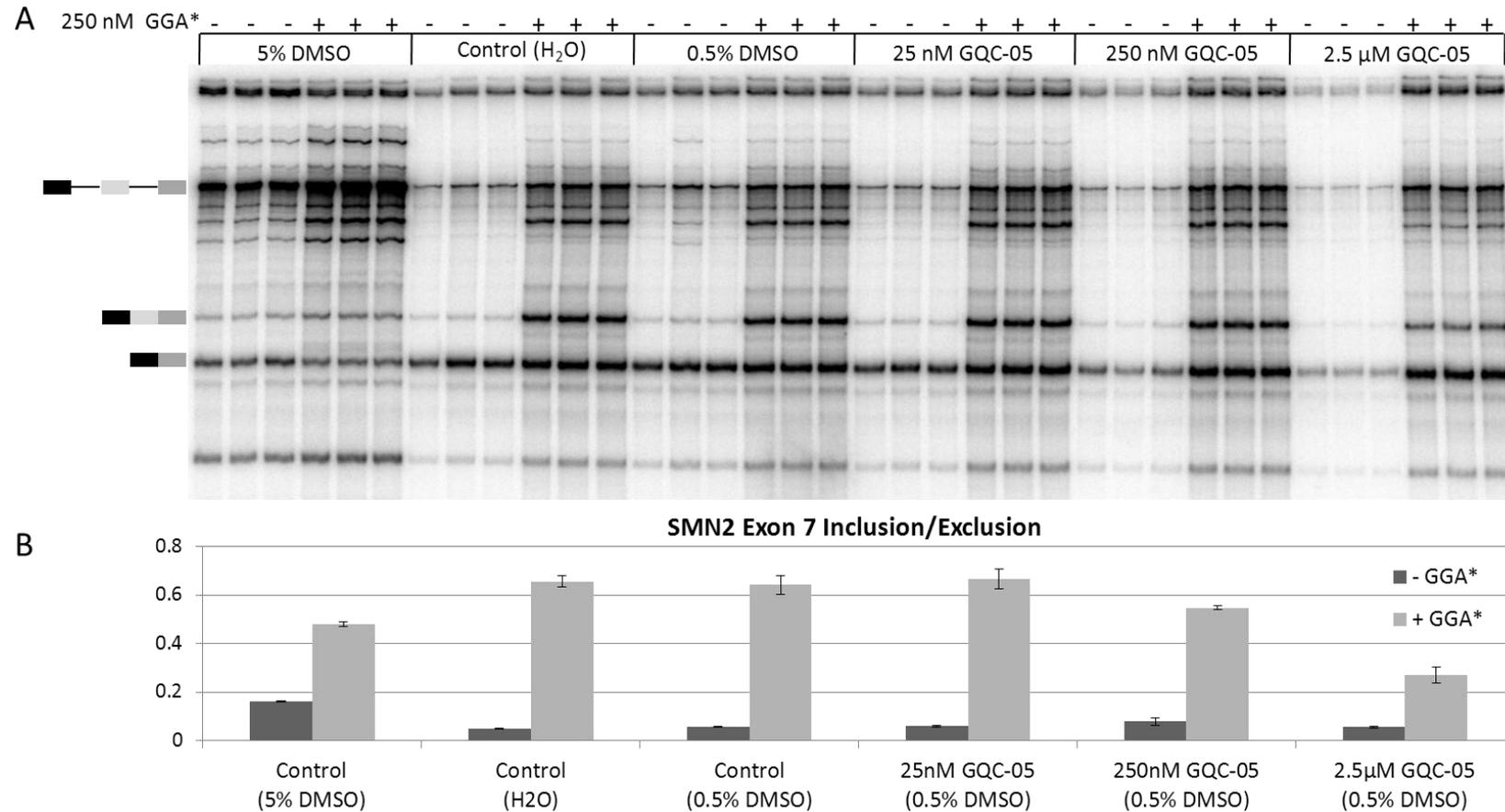


Figure 4.16: (A) Denaturing PAGE of triplicate *in vitro* splicing in HeLa nuclear extract after 2 hrs with radiolabelled SMN2 RNA in the presence of the G-quadruplex stabilizing ligand GQC-05 causes an overall loss of RNA, especially in samples without GGA. This was due to sticking of the RNA to the plasticware used during the experiment, rather than degradation. All further experiments with G-quadruplex stabilizing ligands were done in silanized plasticware to prevent this effect. **(B)** Quantification of the gel in part A (See Section 2.2.7).

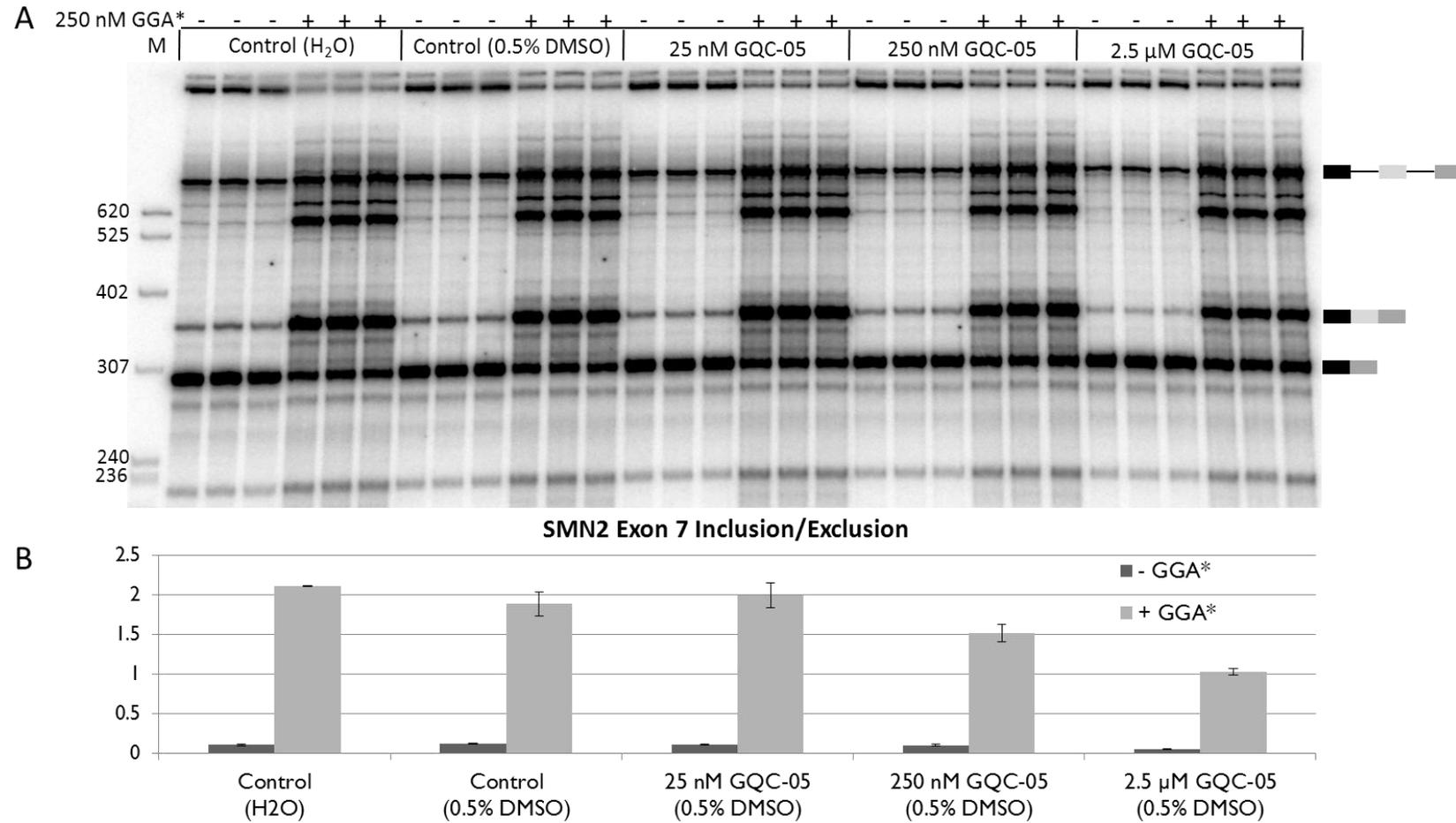


Figure 4.17: GQC-05 reduces SMN2 exon 7 inclusion. (A) Denaturing PAGE of triplicate *in vitro* splicing in HeLa nuclear extract after 2 hrs with radiolabelled SMN2 RNA with and without GGA in the presence of increasing concentrations of GQC-05 G-quadruplex stabilizing ligand. **(B)** Quantification of the gel in part A (See Section 2.2.7).

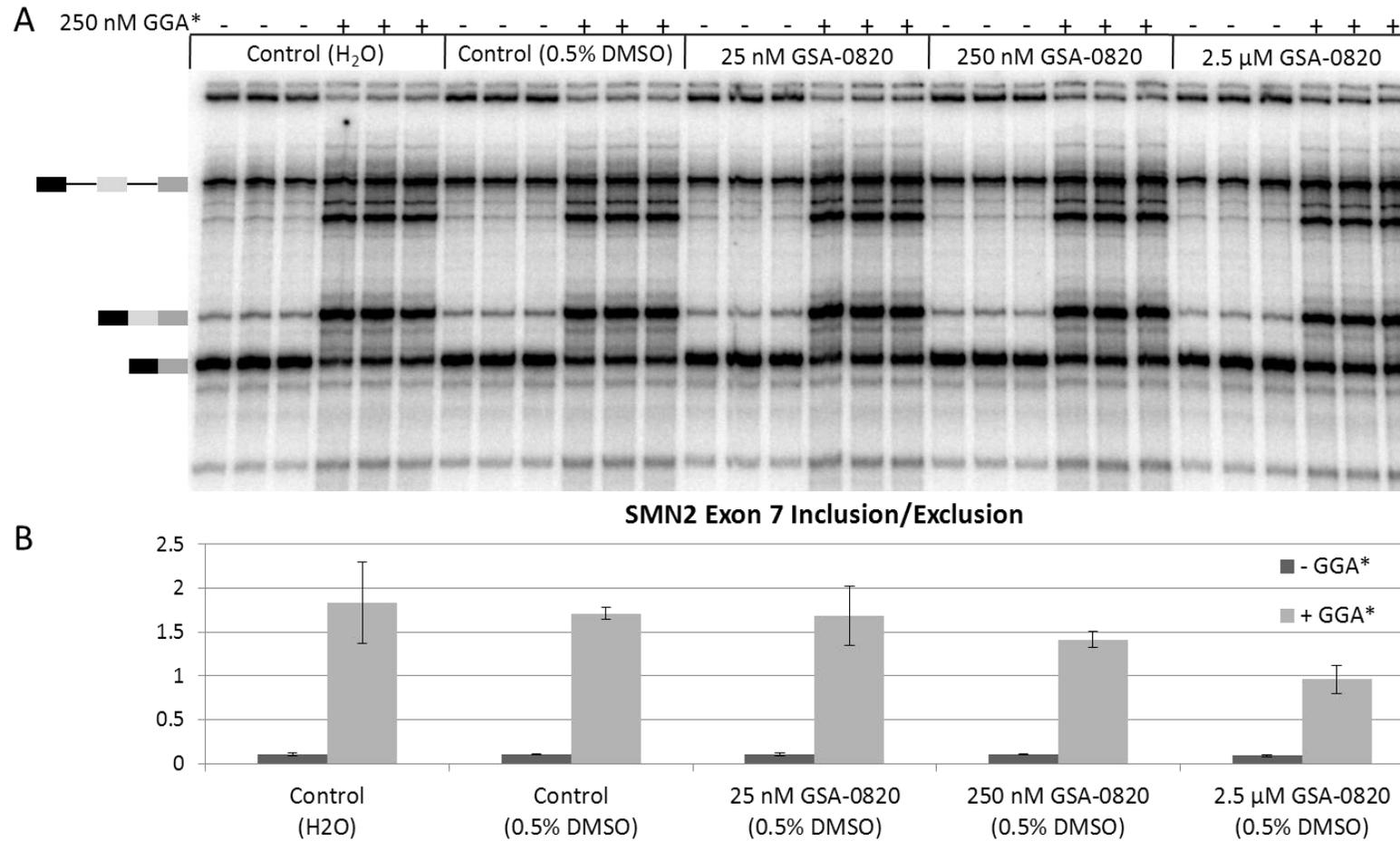


Figure 4.18: GSA-0820 reduces SMN2 exon 7 inclusion. (A) Denaturing PAGE of triplicate *in vitro* splicing in HeLa nuclear extract after 2 hrs with radiolabelled SMN2 RNA with and without GGA in the presence of increasing concentrations of GSA-0820 G-quadruplex stabilizing ligand. **(B)** Quantification of the gel in part A (See Section 2.2.7).

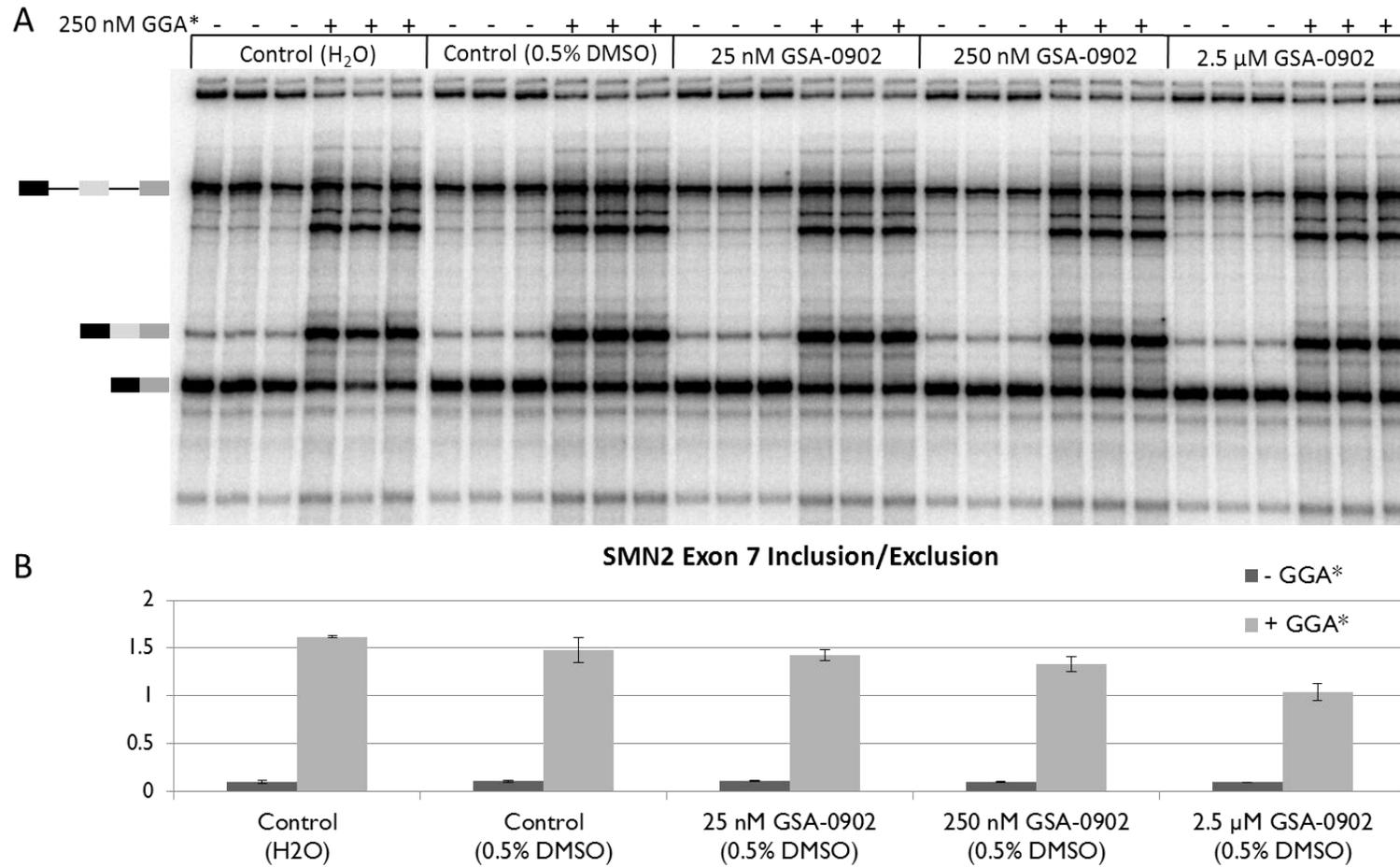


Figure 4.19: GSA-0902 reduces SMN2 exon 7 inclusion. (A) Denaturing PAGE of triplicate *in vitro* splicing in HeLa nuclear extract after 2 hrs with radiolabelled SMN2 RNA with and without GGA in the presence of increasing concentrations of GSA-0902 G-quadruplex stabilizing ligand. **(B)** Quantification of the gel in part A (See Section 2.2.7).

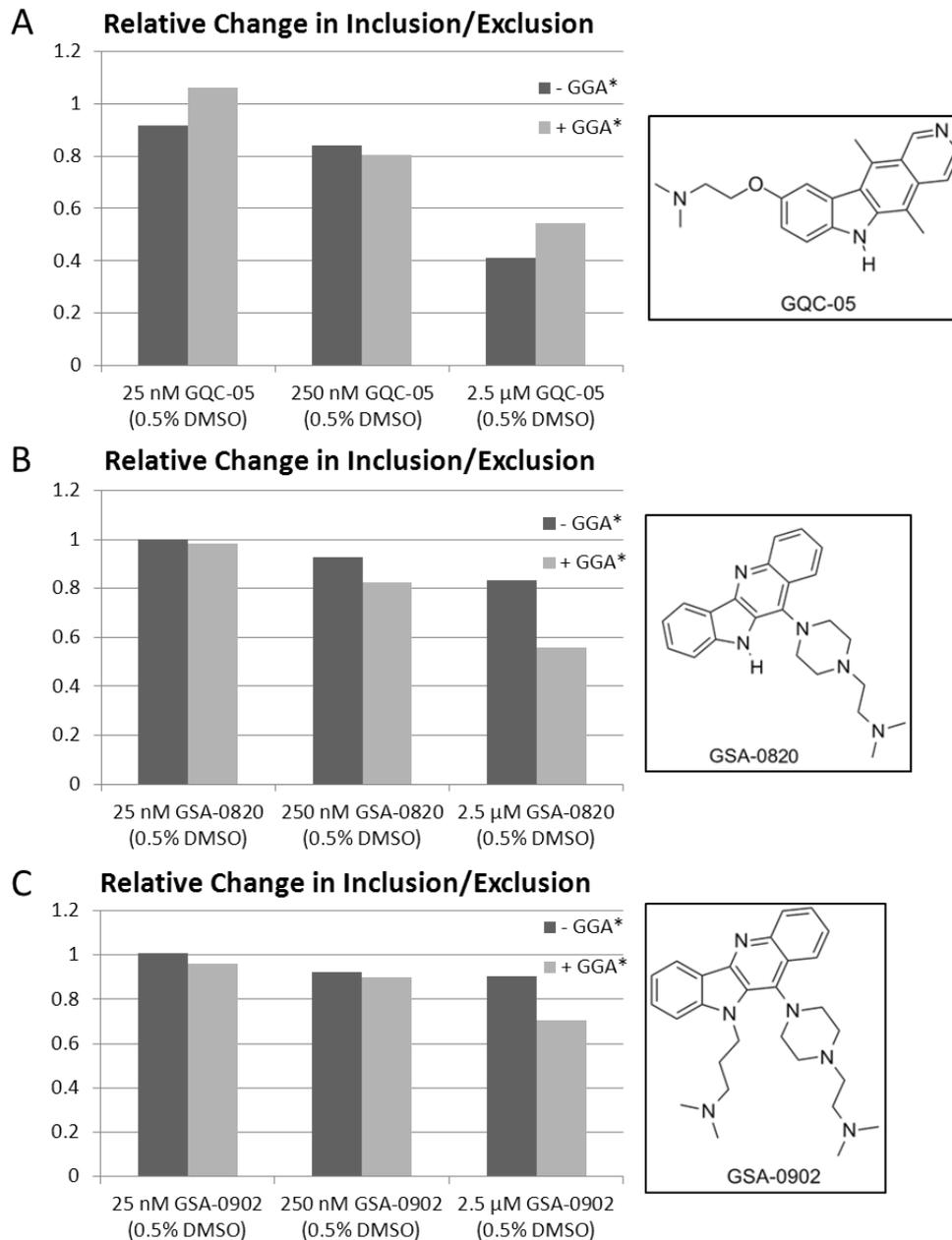


Figure 4.20: GSA-0820 and GSA-0902, both similar in structure, have a greater inhibitory effect on SMN2 exon 7 inclusion in the presence of GGA, indicating a specific effect on the G-quadruplex tail of GGA. However, GQC-05 has a negative effect on SMN2 exon 7 inclusion in the absence of GGA, which is similar or greater than the effect seen in the presence of GGA. **(A)** Graph indicating the relative change in the ratio of SMN2 exon 7 inclusion/exclusion, generated by normalizing the *in vitro* SMN2 splicing data with GQC-05 from Figure 4.17 to the relevant control samples. **(B)** Graph indicating the relative change in the ratio of SMN2 exon 7 inclusion/exclusion, generated by normalizing the *in vitro* SMN2 splicing data with GSA-0820 from Figure 4.18 to the relevant control samples. **(C)** Graph indicating the relative change in the ratio of SMN2 exon 7 inclusion/exclusion, generated by normalizing the *in vitro* SMN2 splicing data with GQC-05 from Figure 4.19 to the relevant control samples.

The next step was to look into the effects of the G-quadruplex stabilizing ligands on protein binding to 5' radiolabelled GGA and TO. For 2.5 μM GQC-05, GSA-0820 and GSA-0902, native polyacrylamide gel electrophoresis was used to assess the effect on complex formation (Figure 4.21A, Figure 4.22A, and Figure 4.23A respectively). In all cases, complex formation was not affected. Interestingly, the GQC-05 samples lacking nuclear extract show a clear upwards shift. This could be evidence of bimolecular G-quadruplex formation. This effect is not seen with GSA-0820 or GSA-0902. Further evidence for a bimolecular G-quadruplex in the presence of 2.5 μM GQC-05 can be found in Figure 4.21B, a UV crosslinking SDS PAGE gel showing crosslinking of GGA to itself in the presence of GQC-05. Again this phenomenon is only observed for GQC-05. No significant changes in protein binding to GGA or TO, as assessed by UV crosslinking, were observed in the presence of 2.5 μM GQC-05, GSA-0820 or GSA-0902 (Figure 4.21B, Figure 4.22B and Figure 4.23B respectively). This suggests that GGA and TO typically exist in G-quadruplex form, as stabilization of this form does not significantly alter protein binding to free GGA and TO. However, this does not explain the negative effects on GGA function observed in the SMN2 *in vitro* splicing experiments.

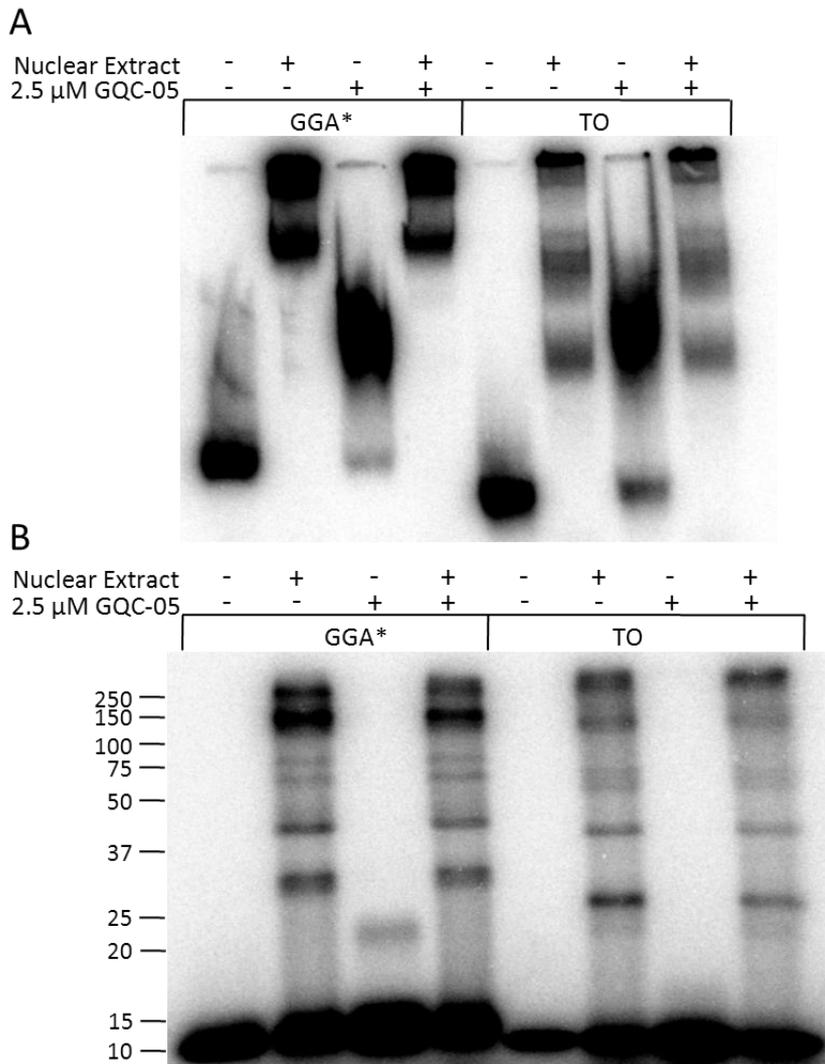


Figure 4.21: Protein binding to GGA and TO is unaffected by 2.5 μ M GQC-05.

(A) Native polyacrylamide gel with 5' radiolabelled GGA and TO in HeLa nuclear extract, showing that in the presence of nuclear extract, complex formation is unchanged by GQC-05. However, without nuclear extract, GQC-05 causes a significant shift in both GGA and TO, possibly indicating the formation of a multi-strand G-quadruplex. **(B)** SDS PAGE gel of UV crosslinking in HeLa nuclear extract with 5' radiolabelled GGA and TO, showing no significant changes to the pattern of protein binding in the presence of GQC-05. Interestingly, for GGA without nuclear extract, a band most likely representing crosslinking of GGA to itself can be seen in the presence of GQC-05, further indicating the potential formation of a multi-stranded G-quadruplex.

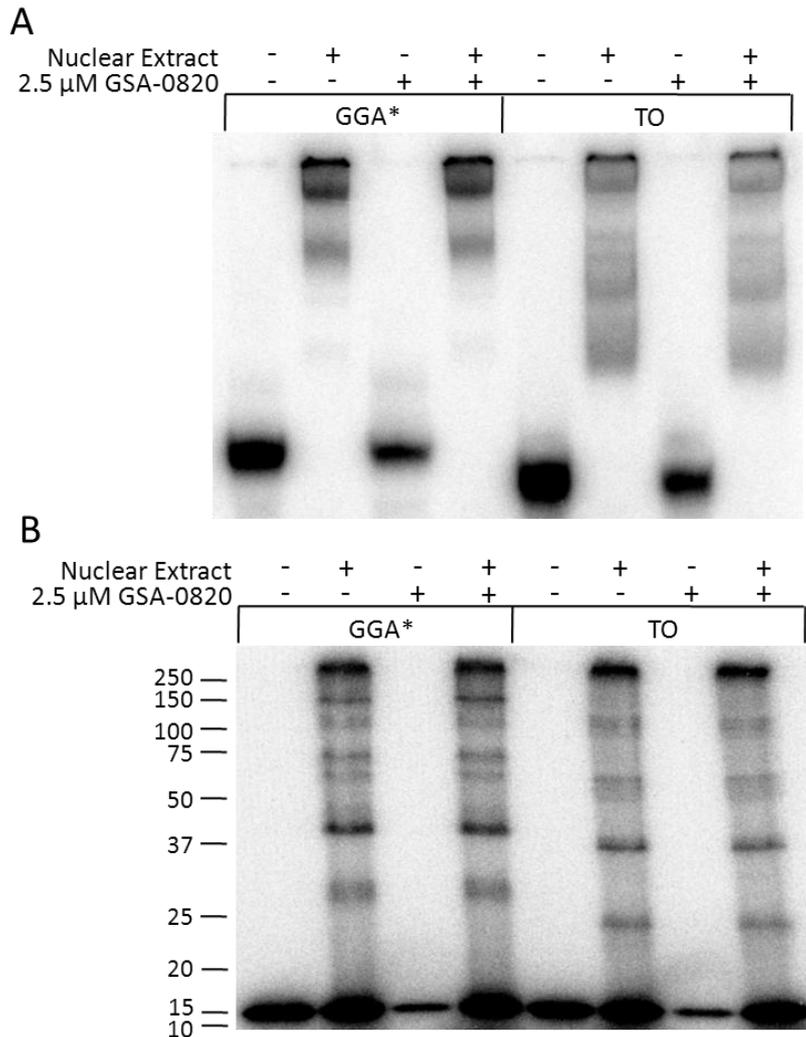


Figure 4.22: Protein binding to GGA and TO is unaffected by 2.5 μ M GSA-0820. **(A)** Native polyacrylamide gel with 5' radiolabelled GGA and TO in HeLa nuclear extract, showing that in the presence of nuclear extract, complex formation is unchanged by GSA-0820. **(B)** SDS PAGE gel of UV crosslinking in HeLa nuclear extract with 5' radiolabelled GGA and TO, showing no significant changes to the pattern of protein binding in the presence of GSA-0820.

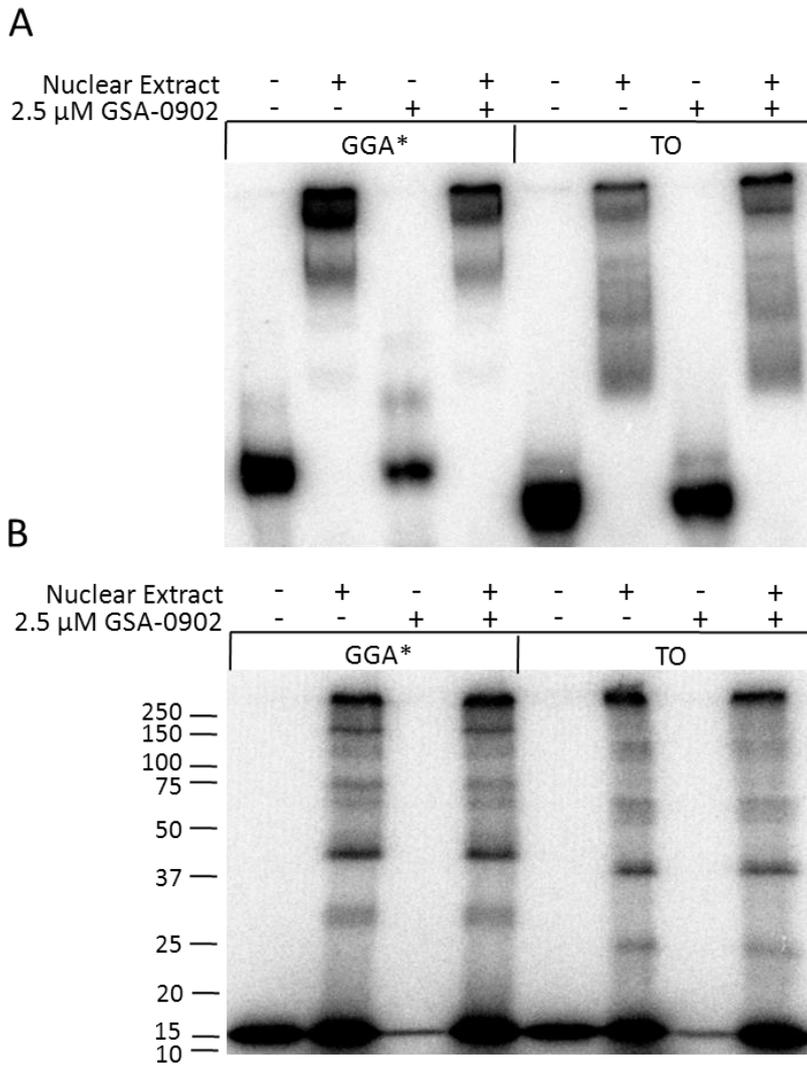


Figure 4.23: Protein binding to GGA and TO is unaffected by 2.5 μ M GSA-0902. **(A)** Native polyacrylamide gel with 5' radiolabelled GGA and TO in HeLa nuclear extract, showing that in the presence of nuclear extract, complex formation is unchanged by GSA-0902. **(B)** SDS PAGE gel of UV crosslinking in HeLa nuclear extract with 5' radiolabelled GGA and TO, showing no significant changes to the pattern of protein binding in the presence of GSA-0902.

Another assay was developed to investigate the G-quadruplex nature of the GGA tail in nuclear extract, reverse transcription along the tail of GGA. Initially the idea was to use SHAPE chemistry (Merino *et al.*, 2005; Wilkinson *et al.*, 2006) to investigate the secondary structure of the tail. However, it soon became evident that the termination pattern from reverse transcription step alone was sufficient to assess the secondary structure. As show in Figure 4.24A, the termination was most prevalent immediately before pairs of G's in the template, especially the first and third pairs of G's, consistent with a G-quadruplex structure. Notably, this pattern is unchanged in the presence of nuclear extract, further indicating that the tail of GGA exists as a G-quadruplex not only in solution, but also in a biologically relevant setting. Figure 4.24B explains that removal of the 5' radiolabel on the primer is why the sample containing nuclear extract in Figure 4.24A is fainter than the control lacking nuclear extract.

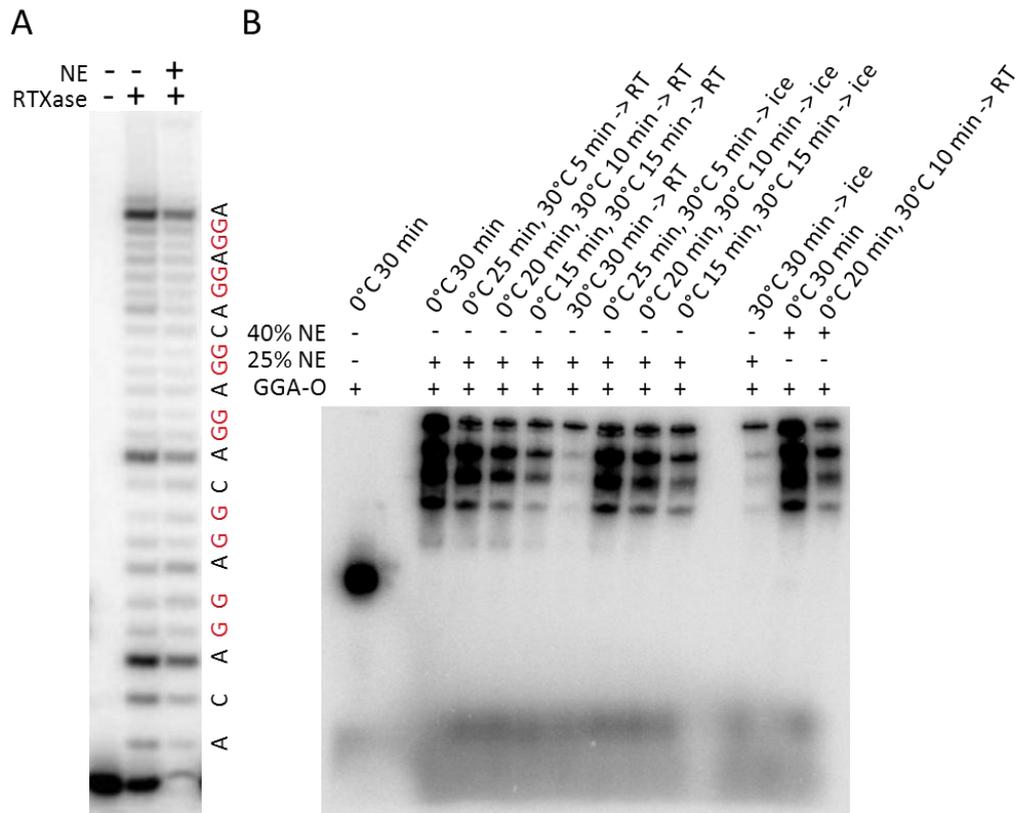


Figure 4.24: Secondary structure analysis of the enhancer tail of GGA is consistent with the formation of a G-quadruplex, even in the presence of nuclear extract.

(A) Reverse transcription along the tail of GGA, using 5' radiolabelled GGA DNA primer (Section 8.2.3) showing termination at the nucleotides just prior to the pairs of G's involved in G-quadruplex formation, as can be seen by the template GGA tail sequence to the right of the image. In the presence of HeLa nuclear extract, the pattern is unchanged. The sample is fainter due to loss of the 5' radiolabel in the presence of nuclear extract, as illustrated in part B. Reverse transcription was done at room temperature. **(B)** Native polyacrylamide gel with 5' radiolabelled GGA-O showing how quickly the 5' radiolabel is removed in the presence of nuclear extract when heated to 30 °C.

4.4 Actions of hnRNP F

Investigations in Section 4.3 were focused around stabilization of the G-quadruplex structure of the GGA tail. In this section, *in vitro* splicing with purified RNA binding domains of hnRNP F, a protein which has been shown to antagonize G-quadruplexes, was used to investigate the G-quadruplex structure of GGA in an opposite manner (Dominguez & Allain, 2006). hnRNP F and hnRNP H, a protein with significant homology to hnRNP F, are known to favor G tracts and were both found in the large scale affinity purification with GGA PC Bio, so it is not inconceivable that they could bind directly to the GGA tail and either melt or block the formation of the G-quadruplex structure (Figure 4.1) (Caputi & Zahler, 2001; Dominguez *et al.*, 2010).

In the first *in vitro* splicing of SMN2, done with increasing concentrations of RRM 1&2 of hnRNP F, the domains seemed to be having a positive effect (Figure 4.25). However, it was later determined that this effect was due entirely to the control buffer differing somehow from the buffer in the stock hnRNP F. The buffer being used initially (20 mM NaPO₄, 50 mM NaCl, 0.1% β-mercaptoethanol, pH 7) was clearly detrimental to splicing, so the hnRNP F RRM 1&2 stock was dialyzed into Buffer D, a potassium based buffer used to store nuclear extract. Unfortunately, this also had a negative effect on splicing because it was still raising the concentration of solutes, like potassium, in the reaction (Figure 4.26). To avoid adding any additional buffer or salt to the splicing reaction, the hnRNP F domains were dialyzed into 100 mM potassium glutamate (K/glu) and 40 mM Hepes pH 7.5 with 0.1% β-mercaptoethanol. The amount of K/glu and Hepes normally added to *in vitro* splicing reactions was reduced to account for the K/glu and Hepes being added with the hnRNP F domains. Therefore, the only addition

to the splicing reaction was the 0.1% β -mercaptoethanol necessary for stabilization of the hnRNP F domains. This strategy was successful in greatly reducing the negative effects on splicing, allowing for the true, very slightly negative, effect of the hnRNP F domains to be visualized (Figure 4.27). Next, hnRNP F RRM3, the strongest of the binding domains, was tested using this method. Again the results showed a negative effect with increasing concentration of hnRNP F (Figure 4.28). To be sure the original positive effect was false, a new batch of hnRNP F RRM 1&2 was prepared in the sodium phosphate buffer. This time the *in vitro* splicing showed the same slightly negative trend as all of the other experiments, except the first (Figure 4.29).

These experiments did not provide any proof in favor of the model where GGA is active in the 'melted' form. However, it is entirely possible that the slight negative effect seen with the addition of the hnRNP F domains was due to competition of the domains with either full length hnRNP F or H, or with an activator protein, such as SRSF1. Therefore, no real conclusions can be drawn one way or the other.

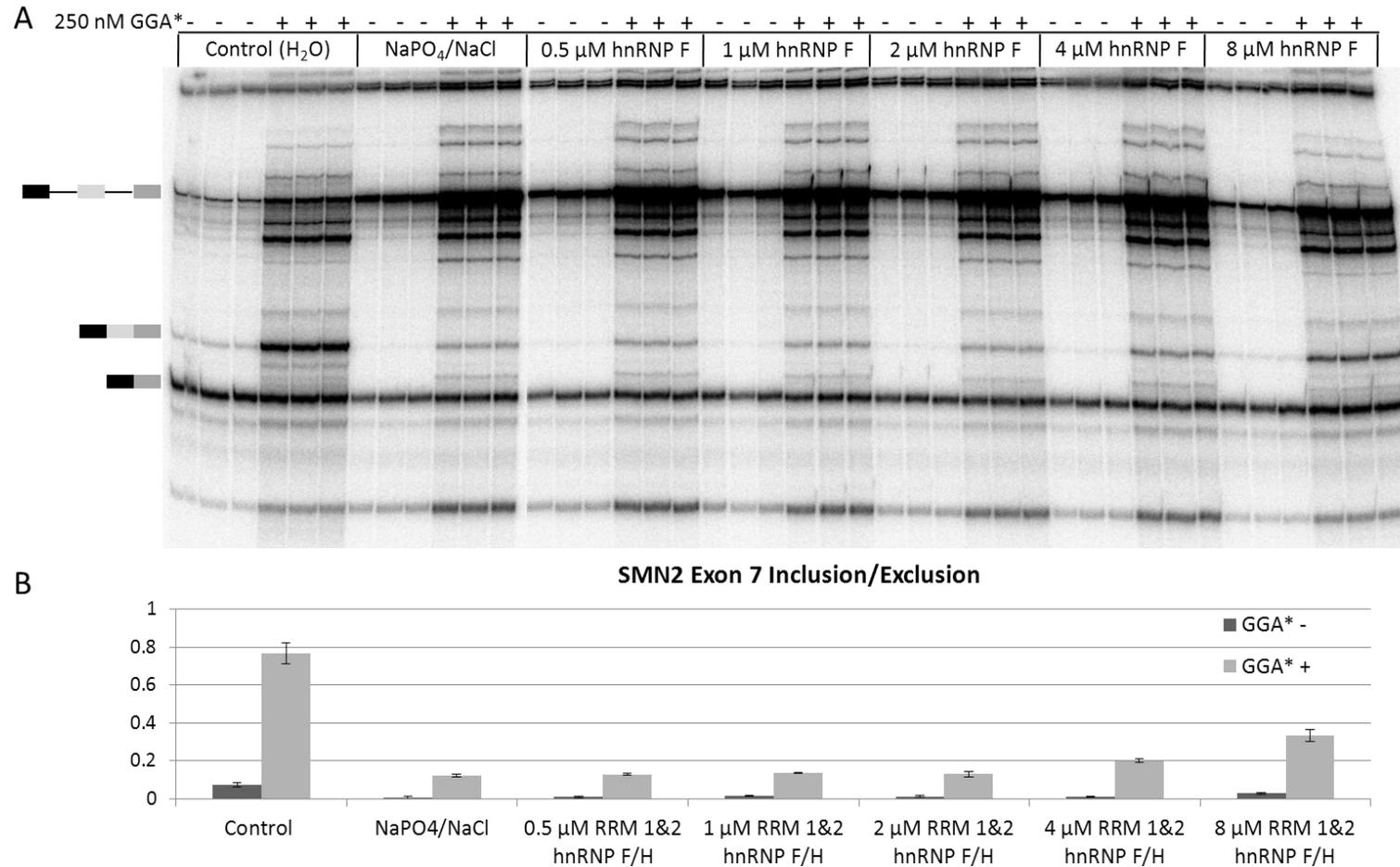


Figure 4.25: (A) Denaturing PAGE of triplicate *in vitro* SMN2 splicing in HeLa nuclear extract with GGA at 2 hrs with purified hnRNP F RNA recognition motifs 1 & 2 showing an apparent increase in exon 7 inclusion. However, this was later shown to be due to a mishap with the control buffer. **(B)** Graph showing the quantification of SMN2 exon 7 inclusion/exclusion from the image in part A (See Section 2.2.7). It is clear that the sodium phosphate buffer has an inhibitory effect on splicing.

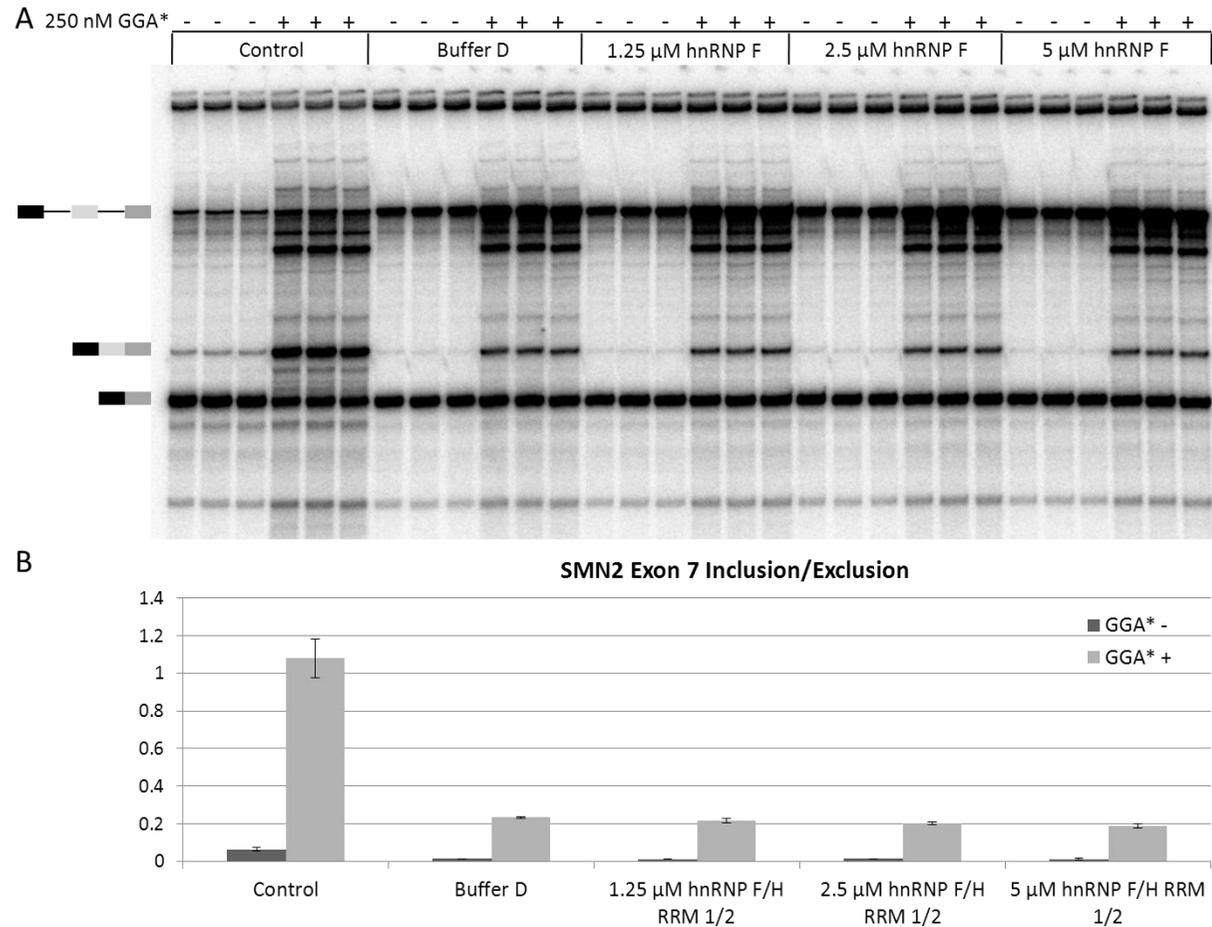


Figure 4.26: (A) Denaturing PAGE of triplicate *in vitro* SMN2 splicing in HeLa nuclear extract with GGA at 2 hrs with purified hnRNP F RRM1 & 2 in Buffer D shows that although all of the components of Buffer D are safe for splicing, the total increase in solutes was still inhibitory. **(B)** Quantification of the gel in part A (See Section 2.2.7) indicates that the hnRNP F domains are not having a positive effect under these conditions. In fact, a slight reduction in exon 7 inclusion is observed.

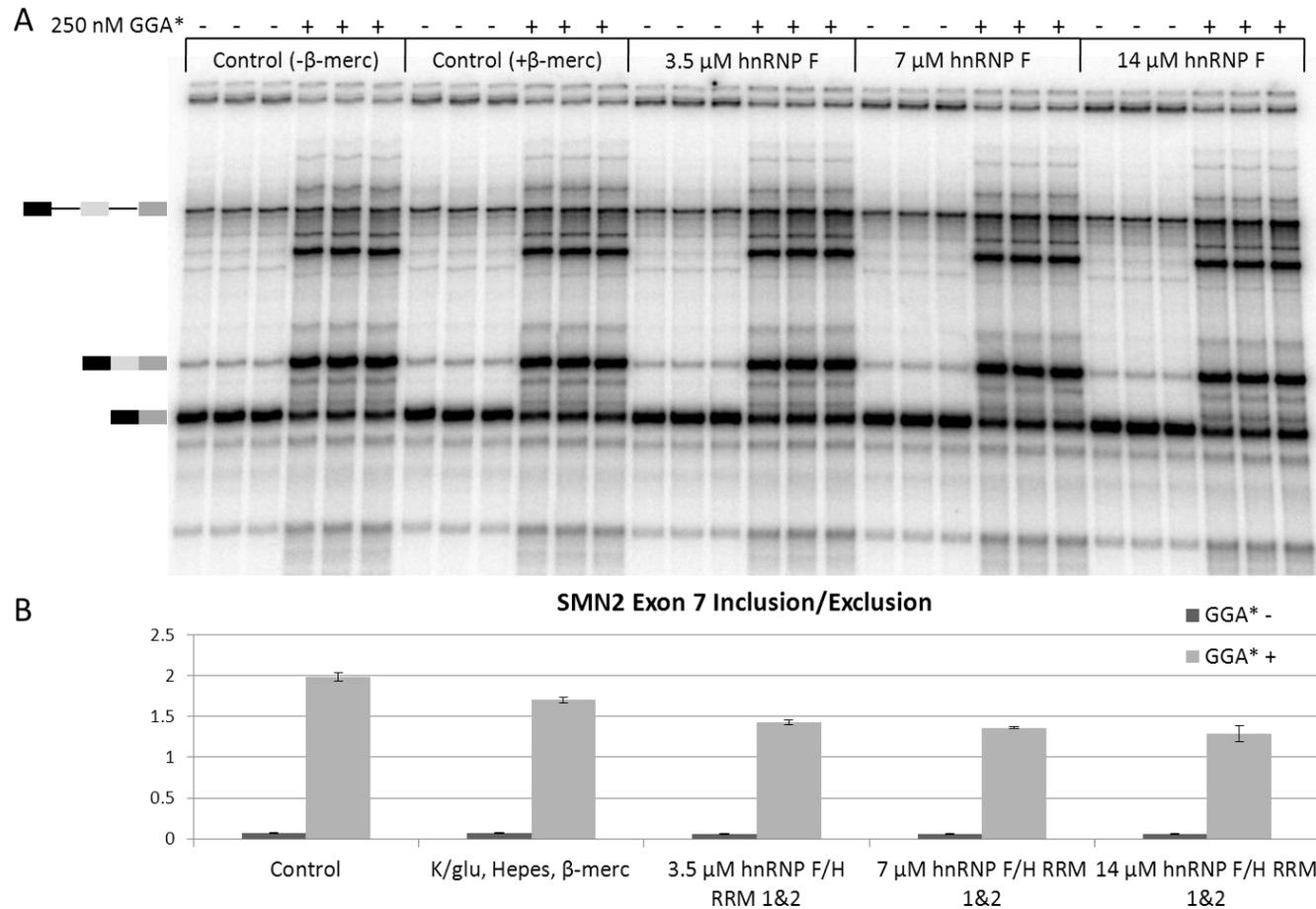


Figure 4.27: (A) Denaturing PAGE of triplicate *in vitro* SMN2 splicing in HeLa nuclear extract with GGA at 2 hrs with purified hnRNP F RRM1 & 2 in K/glu + Hepes buffer, giving final concentrations as for a normal splicing reaction with the addition of 0.1% β-mercaptoethanol. Only very slight inhibition of splicing is observed. **(B)** Quantification of the gel in part A (See Section 2.2.7) shows a small negative effect of hnRNP F RRM1 & 2 on the ratio of SMN2 exon 7 inclusion/exclusion.

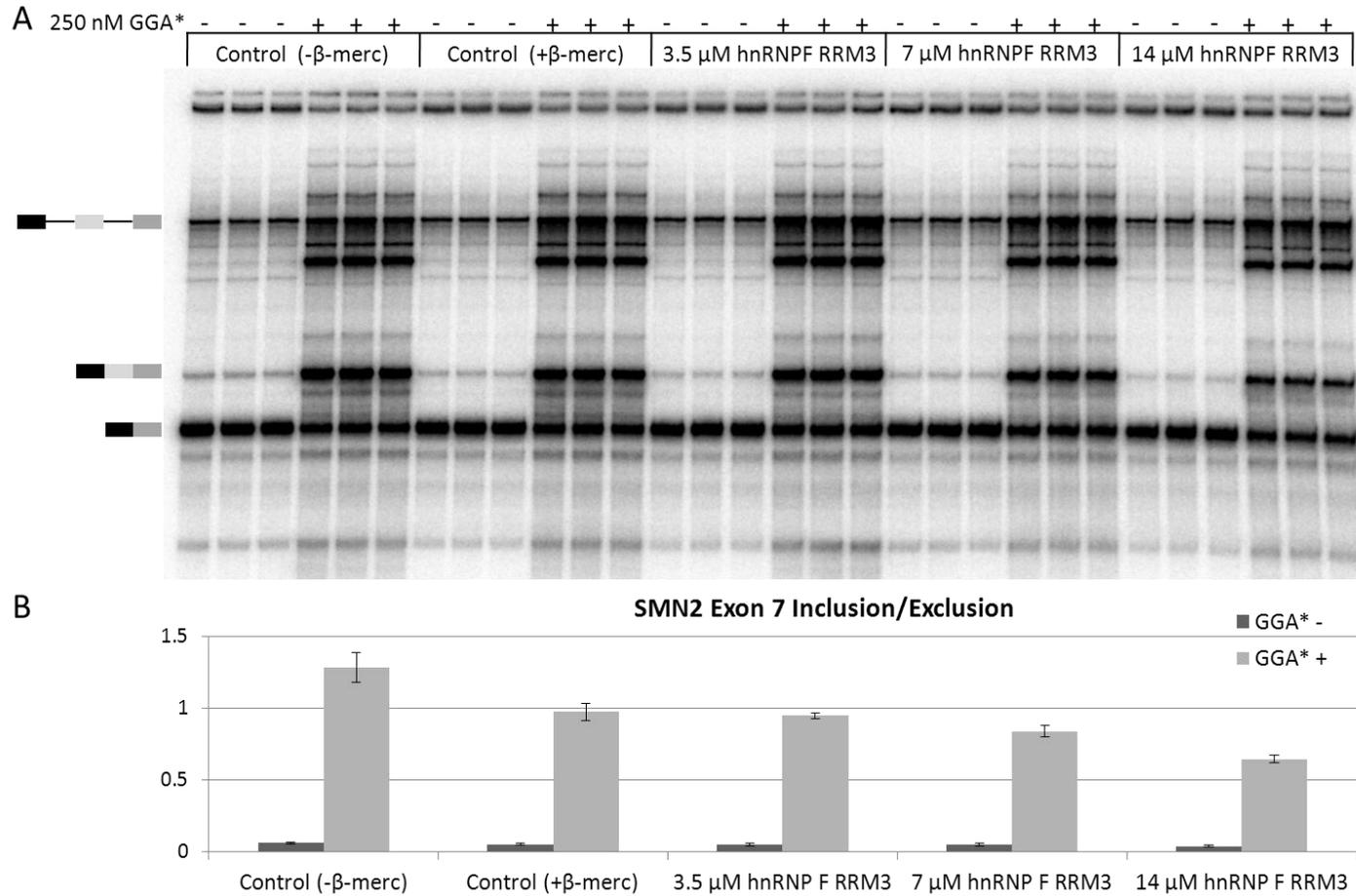


Figure 4.28: (A) Denaturing PAGE of triplicate *in vitro* SMN2 splicing in HeLa nuclear extract with GGA at 2 hrs with purified hnRNP F RRM 3 in K/glu + Hepes buffer, giving final concentrations as for a normal splicing reaction with the addition of 0.1% β-mercaptoethanol. **(B)** Quantification of the gel in part A (See Section 2.2.7) shows a negative effect of hnRNP F RRM3 on the ratio of SMN2 exon 7 inclusion/exclusion.

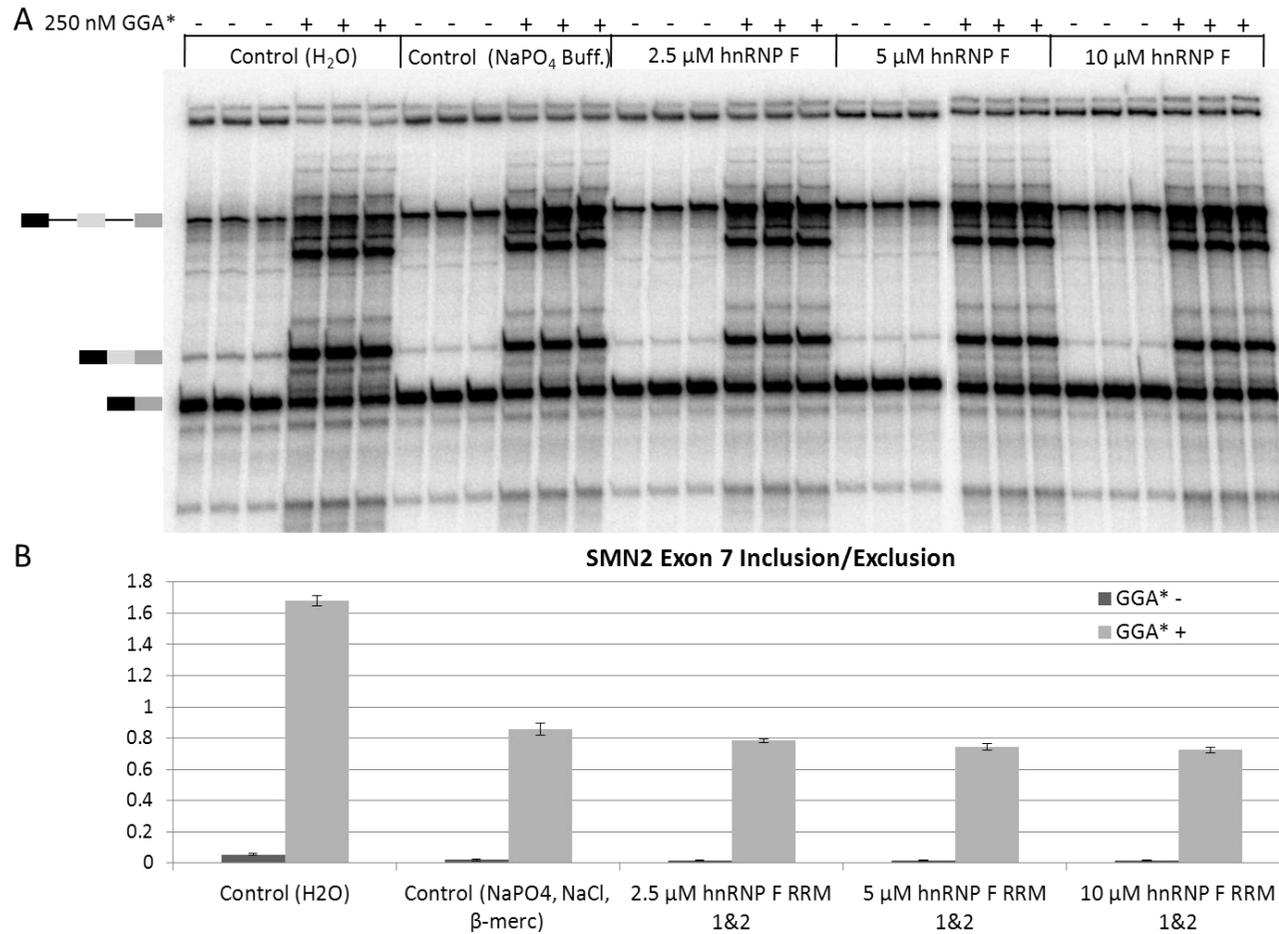


Figure 4.29: (A) Denaturing PAGE of triplicate *in vitro* SMN2 splicing in HeLa nuclear extract with GGA at 2 hrs with a new batch of purified hnRNP F RRM1 & 2 in fresh sodium phosphate buffer further proves that the results from Figure 4.25 were false. **(B)** Quantification of the gel in part A (See Section 2.2.7) shows a slight negative effect of hnRNP F RRM1 & 2 on the ratio of SMN2 exon 7 inclusion/exclusion, as was observed in all related experiments (except for Figure 4.25).

4.5 Importance of Oligonucleotide Chemistry

A study published in 2011 determined, after investigation of many different oligonucleotide sequences and chemistries, that GGA (oligonucleotide 1) was the best of those tested at increasing SMN2 exon 7 inclusion (Owen *et al.*, 2011). Ironically, this is the same sequence and chemistry as was originally used in 2003 when the method was developed (Skordis *et al.*, 2003). This section takes a brief look at a few different oligonucleotides with the same sequence as GGA, but different chemistries, to investigate the effects of oligonucleotide chemistry on protein binding and secondary structure. (See Appendix – Section 8.2.1 for oligonucleotide information).

Figure 4.30 shows that variations in oligonucleotide chemistry affect protein binding and complex formation. UV crosslinking to the 5' radiolabelled oligonucleotides revealed variations in the composition and concentrations of proteins bound (Figure 4.30A). Native polyacrylamide gel electrophoresis of the same samples confirms variations in complex formation (Figure 4.30B). These variations in protein binding are likely due, at least in part, to variations in secondary structure induced by the chemical alterations. Reverse transcription on the tail of various GGA-related oligos indicates variations in secondary structure, consistent with this hypothesis (Figure 4.31). GGA-A, GGA-B, GGA-H, and GGA* are all commercially obtained batches of GGA, which should theoretically be identical in sequence and chemical modifications. Strangely, GGA-A and GGA-B have a different reverse transcription pattern to GGA* despite all three being different batches of the same oligonucleotide with the same modifications. Further investigations of these three batches of GGA, as well as another batch (GGA-H), revealed that loss of phosphorothioate modifications over time could be to blame.

Reverse transcription on the tails of the four batches shows that all are full length (Figure 4.32A). However, high resolution polyacrylamide gel electrophoresis of the oligos shows variations in size, which must therefore be due to loss of phosphorothioate modifications (Figure 4.32B). Figure 4.32C, used with permission from the thesis of colleague Dr L. Smith, shows that loss of phosphorothioate modifications leads to a loss of functionality (Smith, 2012). GGA-A and GGA-B showed the most severe loss of phosphorothioates and likewise showed the least ability to stimulate SMN2 exon 7 inclusion. GGA*, which shows very little to no phosphorothioate loss, performs the best in the SMN2 *in vitro* splicing assay. Most telling perhaps is GGA-H, which has suffered partial phosphorothioate loss and, consequently, partial loss of functionality when compared to GGA*. Clearly, the modifications of GGA* are very important for oligonucleotide secondary structure, protein binding and overall functionality.

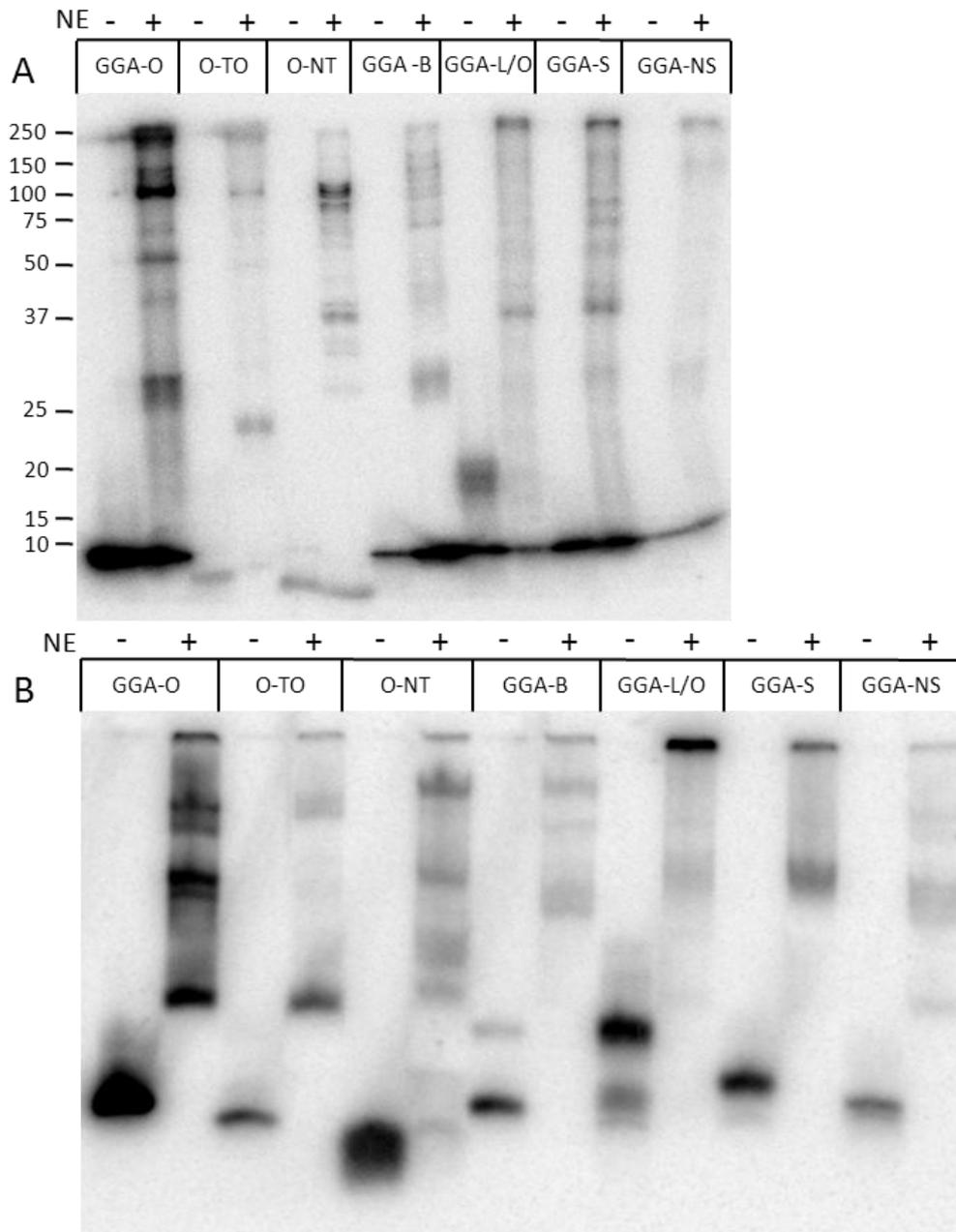


Figure 4.30: Oligonucleotides with the same sequence as GGA, but different chemical modifications (See Section 8.2.1 for detailed oligonucleotide chemical modifications), show significant variations in protein binding. Note: O-TO is the 2'OMe version of the tail-only oligo and O-NT has no tail (annealing region only).

(A) SDS PAGE gel of UV crosslinking with 5' radiolabelled GGA-related oligonucleotides. The various chemistries, such as phosphorothioate (GGA-S), LNA (GGA-L/O), and 2'OMe modifications (GGA-O), cause significant variations in the pattern and identity of proteins bound. **(B)** Native polyacrylamide gel electrophoresis of 5' radiolabelled GGA-related oligonucleotides shows variations in complex formation with the varying chemical modifications.

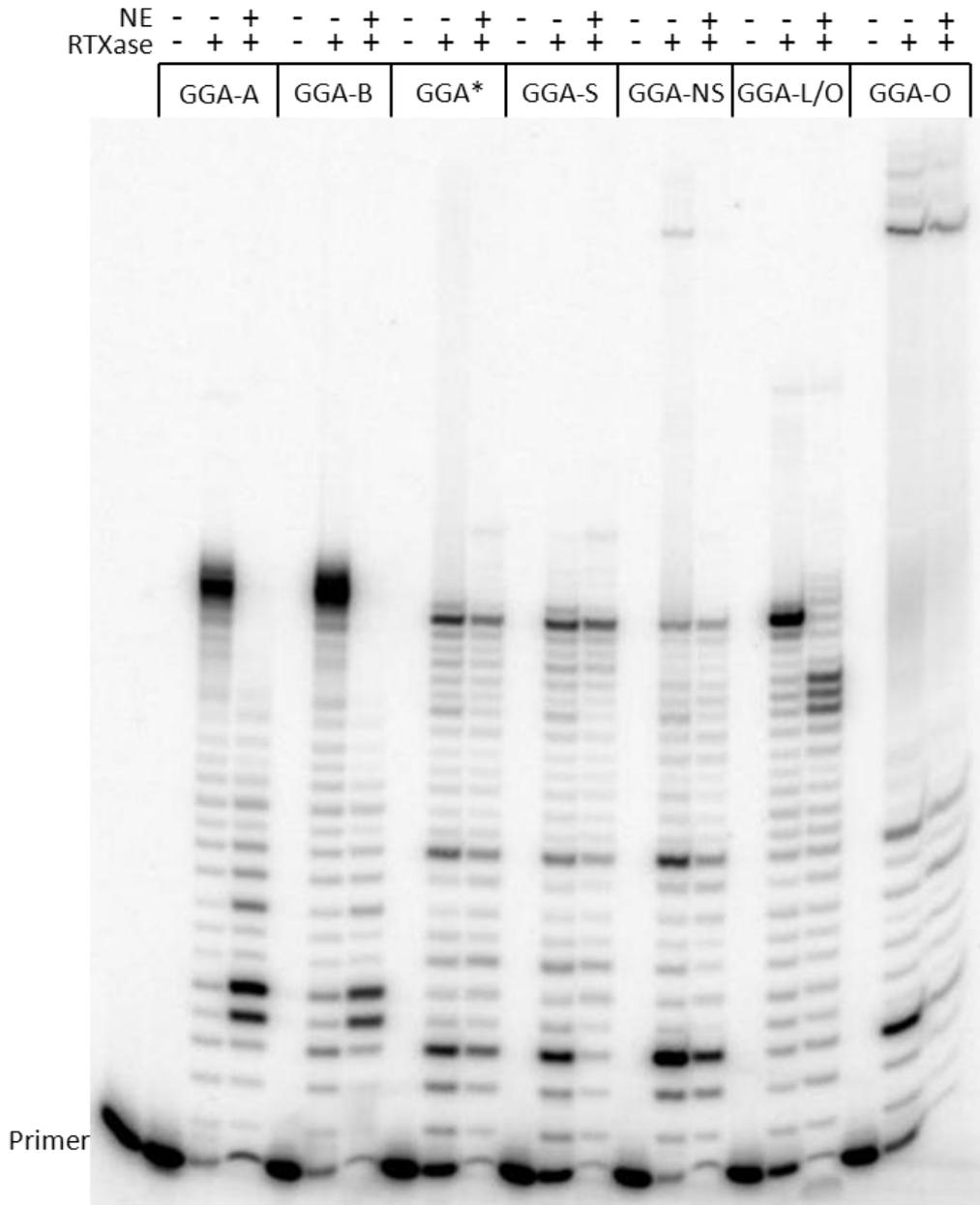


Figure 4.31: Oligonucleotides with the same sequence as GGA, but different chemical modifications (See Section 8.2.1 for details) show variations in secondary structure, as assessed by denaturing PAGE of reverse transcription along the enhancer tails of these oligonucleotides using 5' radiolabelled primer complementary to the 15 nt annealing region of the oligos. Interestingly, different commercial batches of GGA, which should have had exactly the same chemical modifications (GGA-A, GGA-B and GGA*), showed differences in the pattern of termination. GGA-A and GGA-B were different to GGA* and behaved differently in nuclear extract. GGA-O, a 2' OMe version of GGA remained annealed to some of the longer reverse transcription products, obscuring the results.

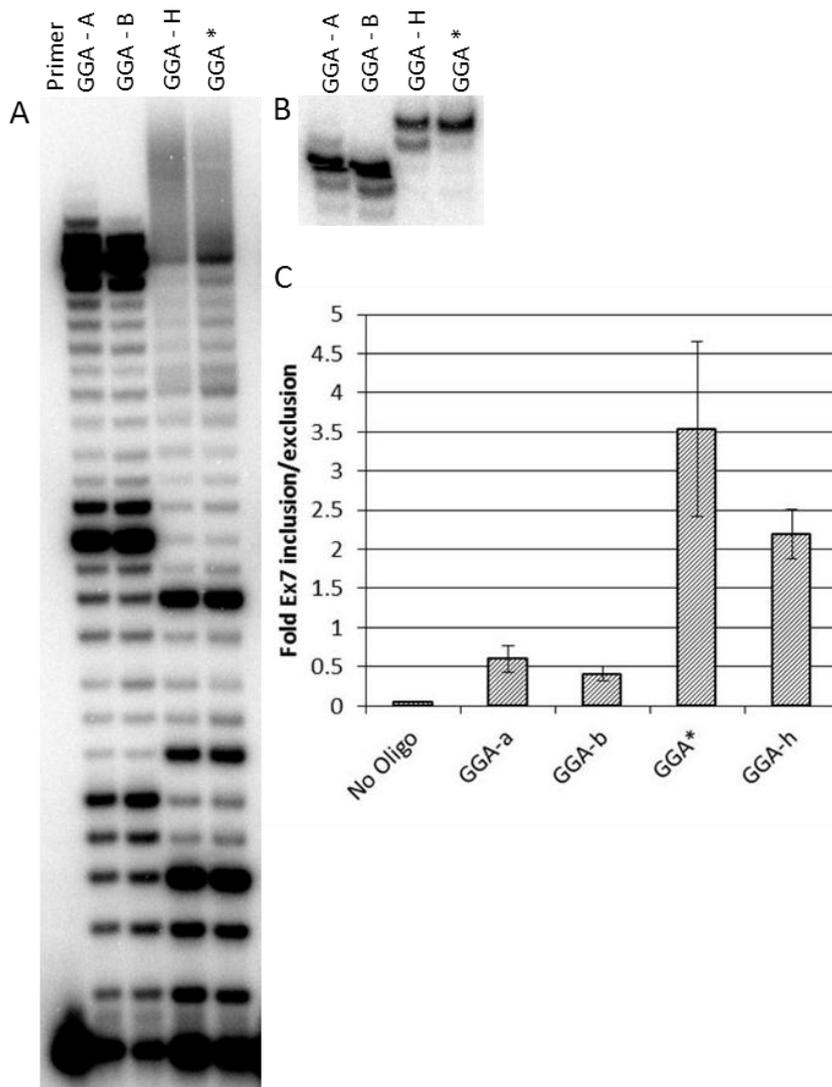


Figure 4.32: The loss of phosphorothioate modifications over time in various batches of GGA (Section 8.2.1) causes changes in secondary structure and a reduction in the ability to stimulate SMN2 exon 7 inclusion.

(A) GGA-A, GGA-B, GGA-H, and GGA* are all commercially obtained batches of GGA, which should theoretically be identical in sequence and chemical modifications. Denaturing PAGE of reverse transcription along the enhancer tails of these oligonucleotides, using a 5' radiolabelled primer annealed to the 15 nt annealing region of the GGA oligos, shows significant variations in secondary structure. **(B)** High-resolution denaturing polyacrylamide gel with 5' radiolabelled batches of GGA, which should have been identical, shows a variation in size. This size difference cannot be due to a loss of nucleotides, as the reverse transcription in part A shows they are all full length. This is therefore most likely due to phosphorothioate modifications reverting back to normal phosphates. **(C)** This graph of the quantification of triplicate *in vitro* SMN2 splicing with the four batches of GGA (performed by Dr L. Smith, adapted with permission from her thesis (Smith, 2012)), shows that the apparent loss of phosphorothioates is nearly proportional to the loss in the ability of GGA to increase the ratio of SMN2 exon 7 inclusion/exclusion.

4.6 Analysis of the Bifunctional Nature of GGA

GGA is a bifunctional oligonucleotide with an annealing region and an enhancer tail, which was designed to be used as a targeted oligonucleotide enhancer of splicing (TOES) for SMN2 exon 7 (Skordis *et al.*, 2003). Although it was not known at the time of design, GGA was found to anneal over a binding site for the splicing repressor hnRNP A1 (Kashima & Manley, 2003; Kashima *et al.*, 2007). Also, a more recent study suggests that there may be a binding site for Sam68 in same region (Pedrotti *et al.*, 2010). Sam68 has been shown to antagonize U2AF65 binding in a different system (Paronetto *et al.*, 2011). Studies of GGA's ability to enhance SMN2 exon 7 inclusion, done by Dr L. Smith (Smith *et al.*, manuscript submitted), indicate that both the annealing region of GGA and the full oligonucleotide are capable of increasing U2AF65 crosslinking to SMN2. However, full length GGA promoted the formation of a U2 dependent complex on exon 7 in addition to increasing U2AF65 binding (Smith, 2012). It was therefore interesting to investigate further the bifunctional nature of GGA.

GGA PC Bio, originally designed for biotin affinity purification experiments, has a biotin on its 3' end and a photocleavable linker between the annealing region and enhancer tail (Figure 3.1). Other than those modifications, it has the same sequence and chemistry as GGA. Triplicate SMN2 splicing comparing GGA with GGA PC Bio revealed that the modifications unique to GGA PC Bio did not largely disrupt its ability to enhance SMN2 exon 7 inclusion (Figure 4.33). The ability to gently cleave the tail of GGA PC Bio using 365 nm UV light made it possible to release the enhancer at various timepoints during SMN2 splicing (Figure 4.34). This revealed a progressive increase in efficacy the longer splicing was allowed to progress with GGA PC Bio before cleavage

of the enhancer tail (Figure 4.34B). Interestingly, even cleavage after 30 minutes caused a nearly 50% reduction in exon 7 inclusion. This implies that the GGA tail may be necessary throughout the entire splicing process. However, further experiments be necessary throughout the entire splicing process. However, further experiments involving cleavage of the tail at specific stalled splicing complexes, followed by a continuation of splicing, would be necessary to be sure.

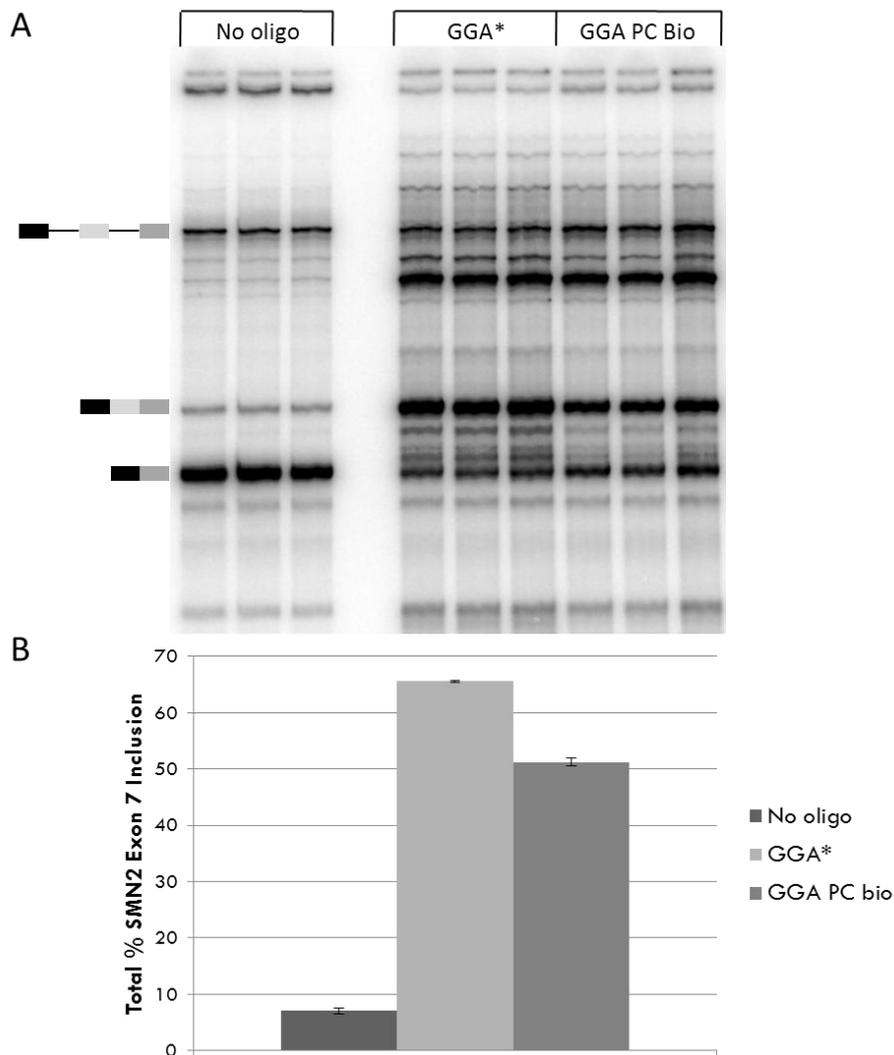


Figure 4.33: GGA PC Bio (Section 8.2.1) is capable of enhancing SMN2 exon 7 inclusion, almost as well as GGA. (A) Denaturing PAGE of triplicate *in vitro* SMN2 splicing in HeLa nuclear extract at 2 hrs with 250 nM GGA and GGA PC Bio. **(B)** Quantification of the percentage of SMN2 exon 7 inclusion of the gel in part A using radiation intensity measurements (See Section 2.2.7).

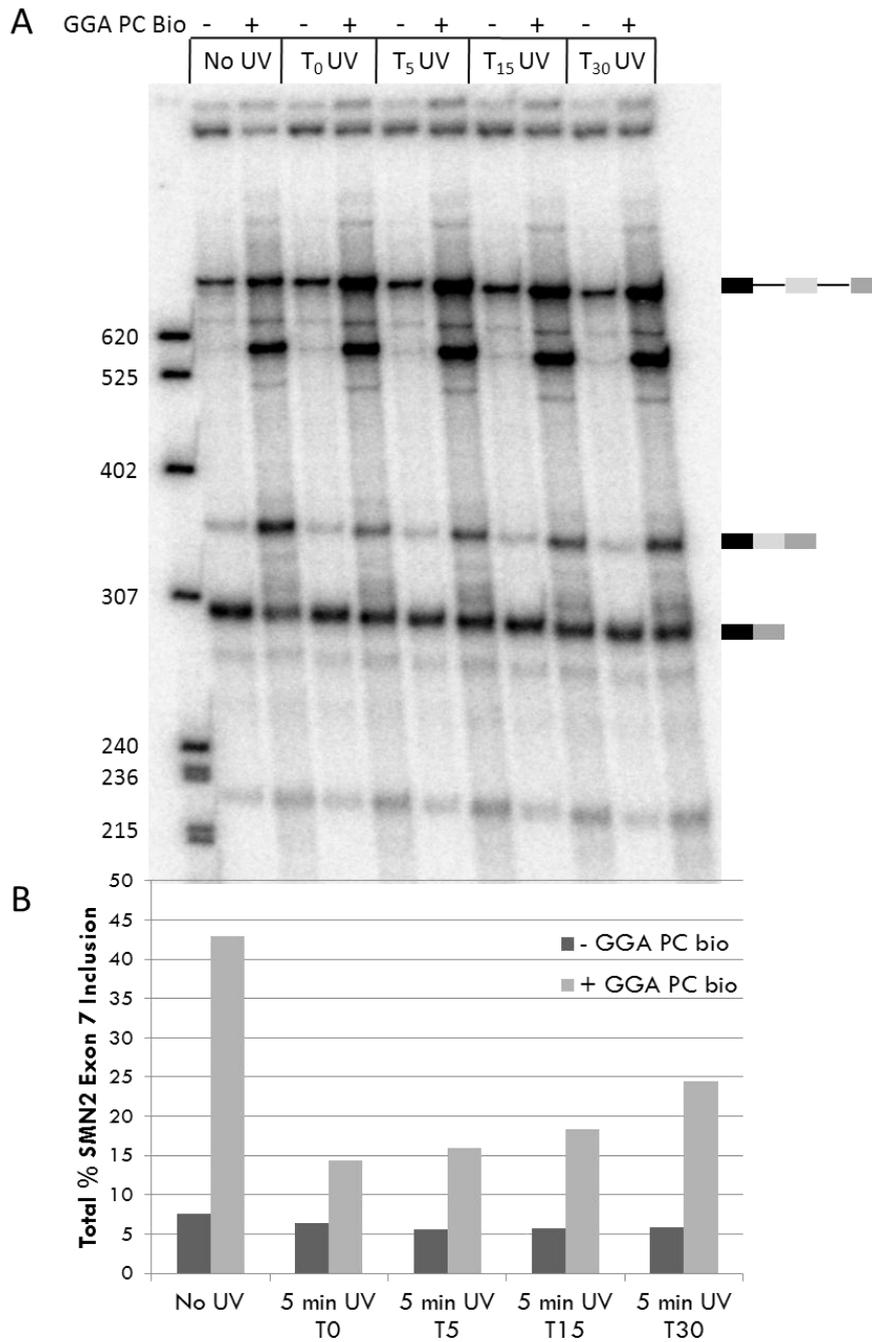


Figure 4.34: The enhancer tail of GGA PC Bio needs to be intact, even after 30 minutes of SMN2 splicing, in order to fully stimulate SMN2 exon 7 inclusion.

(A) Denaturing PAGE of *in vitro* SMN2 splicing in HeLa nuclear extract with GGA PC Bio, where the enhancer tail was removed by photocleavage at various timepoints. **(B)** Quantification of the percentage of SMN2 exon 7 inclusion of the samples from the gel in part A (See Section 2.2.7), which shows that the sooner the enhancer tail of GGA PC Bio is removed, the less of a positive effect it has on exon 7 inclusion.

In order to investigate the potential role of the GGA annealing region in blocking Sam68 binding, single molecule experiments using eGFP-labelled Sam68 and Cy5-labelled SMN2 with and without 250 nM NT were done (Figure 4.35). The single molecule data was analyzed for the overall percent of co-localization of Sam68 with SMN2 as well as the number of labelled Sam68 molecules bound per SMN2 transcript. If the ratio of eGFP-labelled Sam68 to unlabelled Sam68 was known, then the absolute number of Sam68 molecules bound could be determined. However, currently this data has not been obtained for the eGFP-Sam68 nuclear extract that was used. Fortunately, comparison of the results with and without the NT oligonucleotide is still quite informative. Interestingly, the - NT data (Figure 4.35A) shows nearly double the amount of transcripts with 4 or more eGFP-Sam68 molecules bound, as determined by the number of eGFP bleaching steps observed in co-localized spots, compared to the + NT sample (Figure 4.35B). This indicates that the NT oligo is capable of reducing the number of Sam68 molecules bound per transcript. However, the overall number of transcripts bound by at least one Sam68 molecule was increased from 26% to 34% with the addition of 250 nM NT. This may be because the NT oligo reduced aggregation of Sam68, allowing more Sam68 to be available for binding to other SMN2 transcripts. It seems likely from these preliminary results that at least one function of the GGA annealing region is to block Sam68 binding.

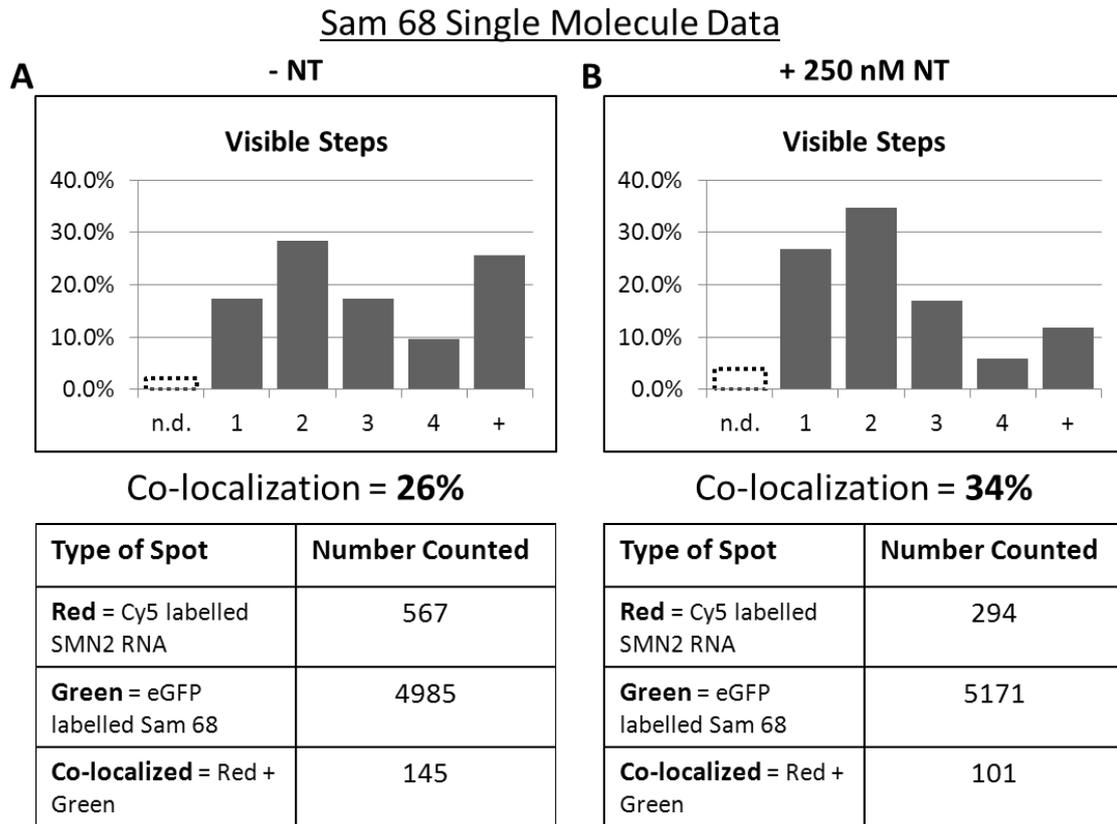


Figure 4.35: The annealing region of GGA (NT) (Section 8.2.1) is capable of reducing the number of Sam68 molecules bound to a single SMN2 transcript, whilst increasing the total number of SMN transcripts with at least one Sam68 bound.

(A) Summary of single molecule data without NT obtained using SMN2 RNA labelled with β -globin 5' Cy5 oligonucleotide (Section 8.2.4) in nuclear extract from HEK 293T cells expressing eGFP-labelled Sam68. The graph of visible steps refers to the number of bleaching steps observed from the eGFP traces. The total percentage of co-localization of labelled Sam68 with labelled SMN2 was 26%, and a large proportion of those co-localized transcripts contained more than 4 Sam68 molecules. **(B)** Summary of single molecule data with 250 nM NT obtained using SMN2 RNA labelled with β -globin 5' Cy5 oligonucleotide (Section 8.2.4) in nuclear extract from HEK 293T cells expressing eGFP-labelled Sam68. The graph of visible steps refers to the number of bleaching steps observed from the eGFP traces. The total percentage of co-localization is increased to 34% (from 26% without NT), but the number of Sam68 proteins bound per transcript is greatly reduced. There is a much higher proportion of transcripts containing one or two labelled Sam68s, and roughly half as many transcripts containing more than 4 Sam68s when compared to the – NT experiment in part A.

4.7 Discussion

Investigation into the mechanism of action of GGA, a bifunctional oligonucleotide that stimulates the inclusion of exon 7 during SMN2 pre-mRNA splicing, led to several interesting findings. First of all, the 23 nt enhancer tail domain (AGGAGGACGGAGGACGGAGGACA) was found to bind the splicing activator protein SRSF1, as expected, as well as the G-quadruplex associated proteins CNBP and nucleolin. This G-rich tail domain has been shown by CD and NMR studies to form a G-quadruplex structure in solution (Mr A. Cousins and Prof. M. Searle, University of Nottingham, (Smith *et al.*, manuscript submitted)). The spontaneous termination pattern seen during reverse transcription along the tail of GGA indicates that this structure is present in nuclear extract as well. This provides some explanation for the association of CNBP and nucleolin, neither of which has been shown to be involved in splicing activation. G-quadruplexes have been shown to be involved in many aspects of RNA biology, including pre-mRNA splicing (Millevoi *et al.*, 2012), so it was important to assess what effect this structure has on the function of GGA. Stabilization of the G-quadruplex structure with ligands resulted in decreased efficacy of GGA in *in vitro* SMN2 splicing assays, indicating that the G-quadruplex most likely melts when bound by the 'active' protein, which seems to be SRSF1. Biotin affinity purification with GGA PC bio, a 3' biotinylated version of GGA with a photocleavable linker between the annealing and enhancer domains, revealed that the GGA tail domain has the ability to recruit a large number of splicing and transcriptional proteins. This is most likely through protein-protein interactions that ultimately result in the association of GGA with pre-mRNAs in the nuclear extract. The enhancer tail domain is vital to the

function of GGA, and release of the tail via photocleavage during a splicing reaction results in greatly reduced functionality. The annealing region of GGA also plays an important role of blocking a splicing silencer sequence, which has been shown to bind hnRNP A1 and/or Sam68 (Kashima & Manley, 2003; Kashima *et al.*, 2007; Pedrotti *et al.*, 2010). Single molecule studies with eGFP-labelled Sam68 and Cy5-labelled SMN2 RNA show that the GGA annealing domain is capable of reducing the number of Sam68 proteins bound per transcript. The annealing and enhancer domains of GGA each have important mechanistic roles, which combine to produce a very potent stimulator of SMN2 exon 7 splicing. Another important finding relating to the efficacy of GGA is that the phosphorothioate modifications of GGA are very important for achieving a high level of SMN2 exon 7 inclusion and that they can revert back to normal phosphate groups over time. This is important because phosphorothioate modifications are used in many potential therapeutic oligonucleotides, and the loss of phosphorothioates in these oligos may result in a similar reduction in efficacy as seen for GGA. All of these findings regarding the structure and function of GGA, in addition to sparking further interest into the role of G-quadruplexes in pre-mRNA splicing, provide important information for the design of future splice-altering oligonucleotides.

5 NEW STRATEGIES FOR RESCUING THE SPLICING OF SMN2 EXON 7 USING OLIGONUCLEOTIDES

5.1 Introduction

5.2 Use of ESE Conjugated Gold Nanoparticles to Enhance SMN2 Exon 7 Inclusion

5.3 *In vitro* Splicing with Morpholino Oligonucleotides Targeting a Silencer in SMN2 Intron 7

5.4 Discussion

5.1 Introduction

Antisense oligonucleotide based therapeutic strategies for the treatment of spinal muscular atrophy, via promotion of SMN2 exon 7 splicing, are very promising (Section 1.3.2). These oligonucleotides alter splicing either by blocking the binding of repressor proteins, recruiting activator proteins, or both. There is still much to learn about how best to design and utilize these types of antisense oligonucleotides. For example, oligonucleotide chemistry has very significant effects on the efficacy and toxicity of antisense oligonucleotides. Therefore, efforts to improve the current strategies or create new strategies are needed in order to improve the design and therapeutic potential of future antisense oligonucleotide based therapies.

5.2 Use of ESE Conjugated Gold Nanoparticles to Enhance SMN2 Exon 7 Inclusion

In an effort to better understand, and possibly improve, the current splicing enhancement strategies using tethered RNA enhancers, we investigated the effects of accessibility and valency on tethered enhancer function (Perrett *et al.*, 2013). This section focuses on the question of valency, which was investigated using gold nanoparticles (GNPs) capable of conjugating many RNA oligonucleotide strands. Several sizes of nanoparticles from 20 nm down to 5 nm were tested. It was found that the 5 nm GNPs were best with the larger sizes causing increasing amounts of splicing inhibition, especially at higher concentrations. This section looks at *in vitro* SMN2 splicing with three 5 nm GNP conjugates, as well as some unpublished supporting work.

GNP 16 (Figure 5.2). It is important to note that while the overall concentration of nanoparticles was the same as the concentration of control oligonucleotide used, each GNP (16 and 17) was conjugated with over 400 strands of RNA (Perrett *et al.*, 2013). Nuclear extract was dialyzed to remove the DTT, preventing reduction of the disulfide bonds conjugating the RNA strands to the nanoparticles. Interestingly, at 50 nM and 100 nM concentrations, GNP 16, containing only the ESE sequence, gave a higher total level of SMN2 exon 7 inclusion than O-TO and GGA-O, while GNP 17, containing the ESE and annealing region, had a negative effect (Figure 5.3A). At the higher 200 nM concentration, GNP 16 still increased the ratio of exon 7 inclusion to exclusion, but the overall level of inclusion was reduced due to a reduction in splicing efficiency (Figure 5.3). This result was unexpected, as GNP 16 lacks the annealing region required to target exon 7. These findings suggest that GNP 16 acts somehow as a non-specific enhancer, which is more powerful, at least at low concentrations, than a targeted bifunctional oligonucleotide with the same enhancer sequence.

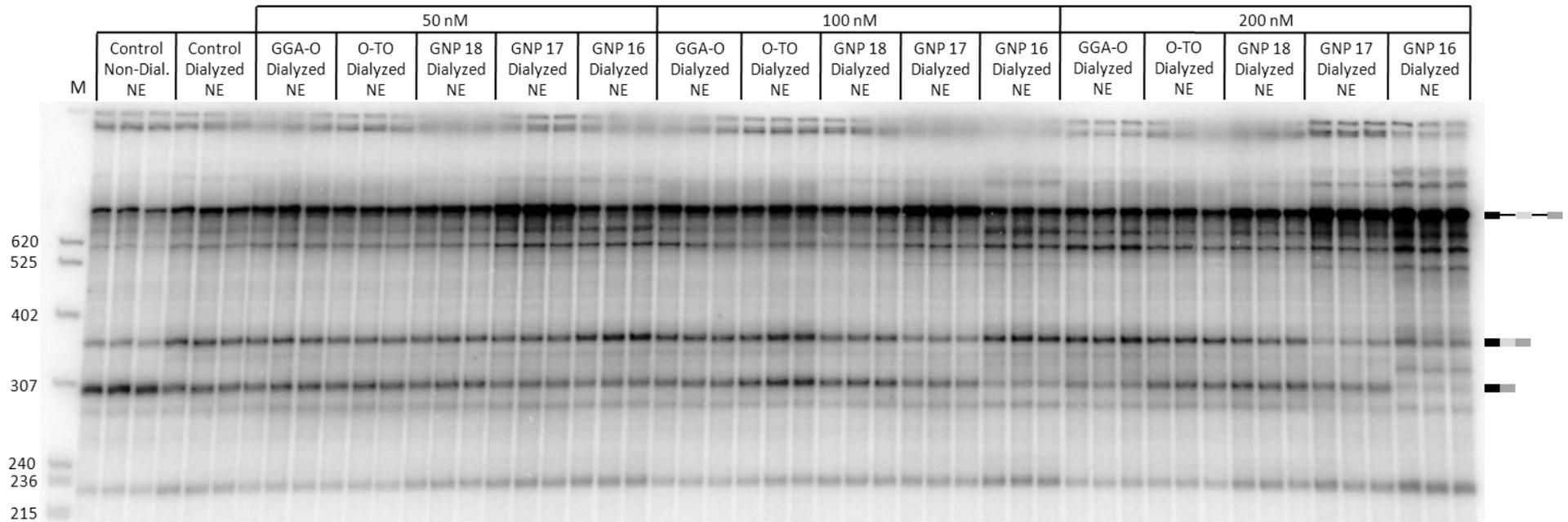


Figure 5.2: Denaturing PAGE of triplicate *in vitro* SMN2 splicing in HeLa nuclear extract at 2 hrs with increasing concentrations of gold nanoparticles modified with GGA related oligonucleotides and control oligonucleotides, as described in Figure 5.1. Quantification (See Section 2.2.7) and analysis of these results can be seen in Figure 5.3.

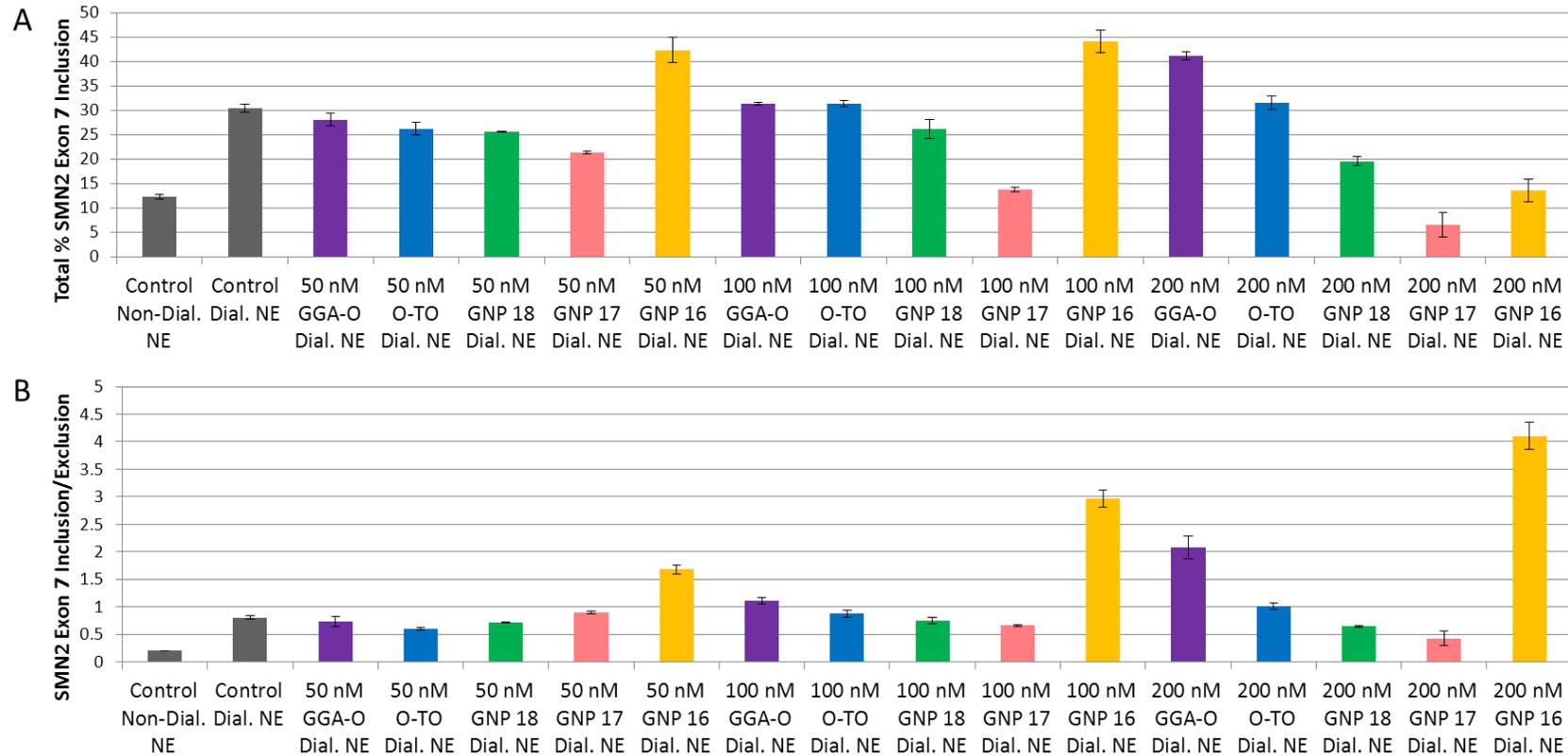


Figure 5.3: Quantification of the SMN2 *in vitro* splicing results with modified gold nanoparticles from Figure 5.2.

(A) Graph showing the total percentage of SMN2 exon 7 inclusion. Only GNP 16, modified with the GGA enhancer tail sequence, showed a greater level of exon 7 inclusion than GGA-O at 50 nM and 100 nM. However, at 200 nM the GNPs had an inhibitory effect on splicing, resulting in a low overall level of exon 7 inclusion. **(B)** Graph showing the ratio of SMN2 exon 7 inclusion/exclusion.

To investigate the apparent ability of GNP 16 to act as a non-specific enhancer, *in vitro* splicing assays were done with single-intron adenovirus constructs containing two alternative consensus 5' splice sites, which would be sensitive to the presence of a non-specific enhancer (Sections 8.1.5 and 8.1.6, provided by Dr H. Lewis, University of Leicester). If GNP 16 was acting as a non-specific enhancer, a shift to the intron-proximal (downstream) 5' splice site would be expected. Triplicate A2 splicing with GNP 16 and relevant controls showed a shift towards the use of the downstream 5' splice site, which increased with increasing concentration of GNP 16 up to 100 nM (Figure 5.4 and Figure 5.5A). However, the overall level of splicing decreased with increasing concentration of GNP 16, and the decline was rather severe at the highest tested concentration of 200 nM (Figure 5.5B). Triplicate splicing with GNP 16 was also done using A3, a construct containing the enhancer sequence of the GGA tail at its 5' end (Figure 5.6). This additional enhancer sequence shifts the 5' splice site selection in favor of the upstream splice site (Lewis *et al.*, 2012). However, increasing concentrations of GNP 16 progressively overcame the effect of this 5' ESE and shifted selection towards the downstream splice site, as was seen for the A2 construct (Figure 5.7A). Again, high concentrations of GNP 16 were inhibitory (Figure 5.7B). These results suggest that GNP 16 is able to somehow non-specifically associate to pre-mRNA, where it functions as an ESE.

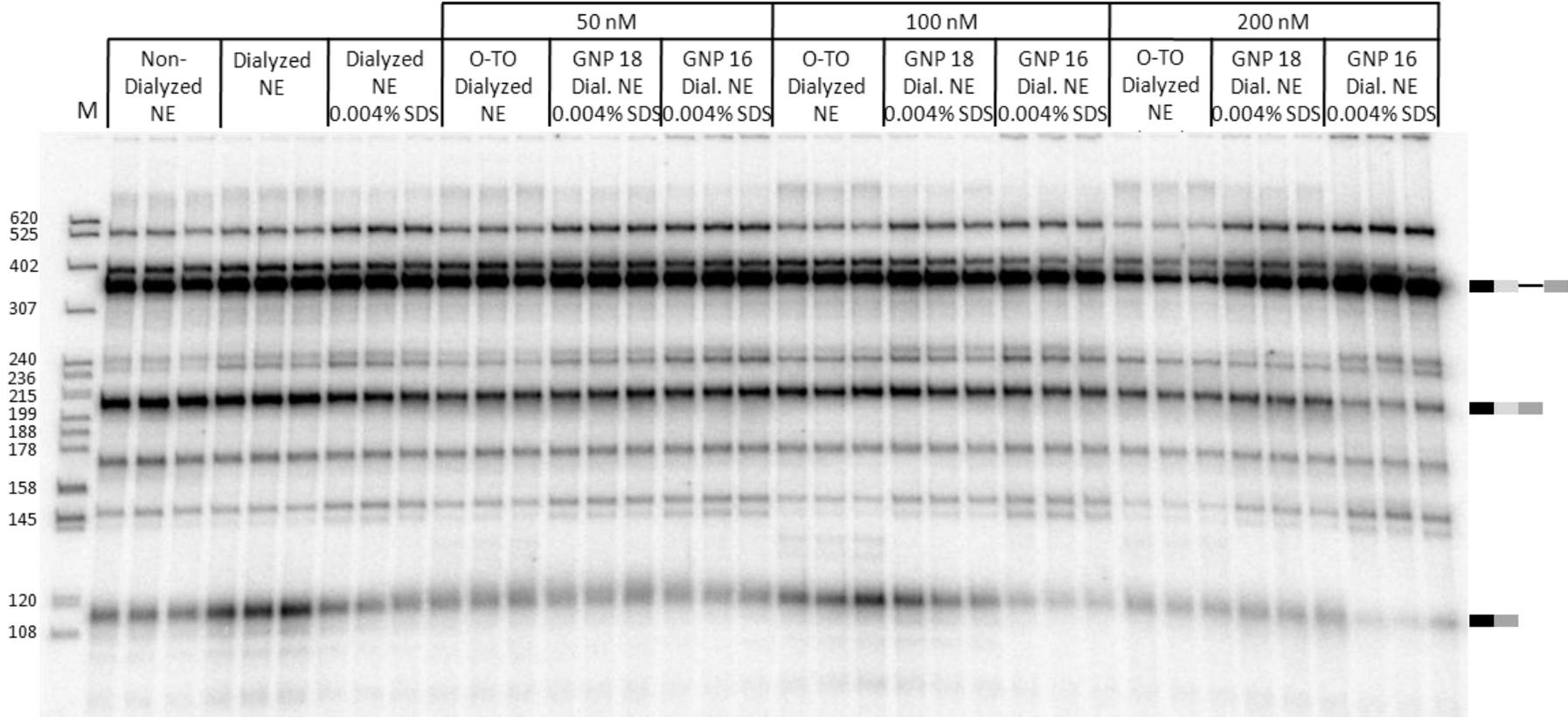


Figure 5.4: Denaturing PAGE of triplicate *in vitro* splicing of adenovirus-derived transcript A2 in HeLa nuclear extract at 2 hrs with increasing concentrations of GNP 16 (Figure 5.1). Quantification (See Section 2.2.7) and analysis of these results can be seen in Figure 5.5.

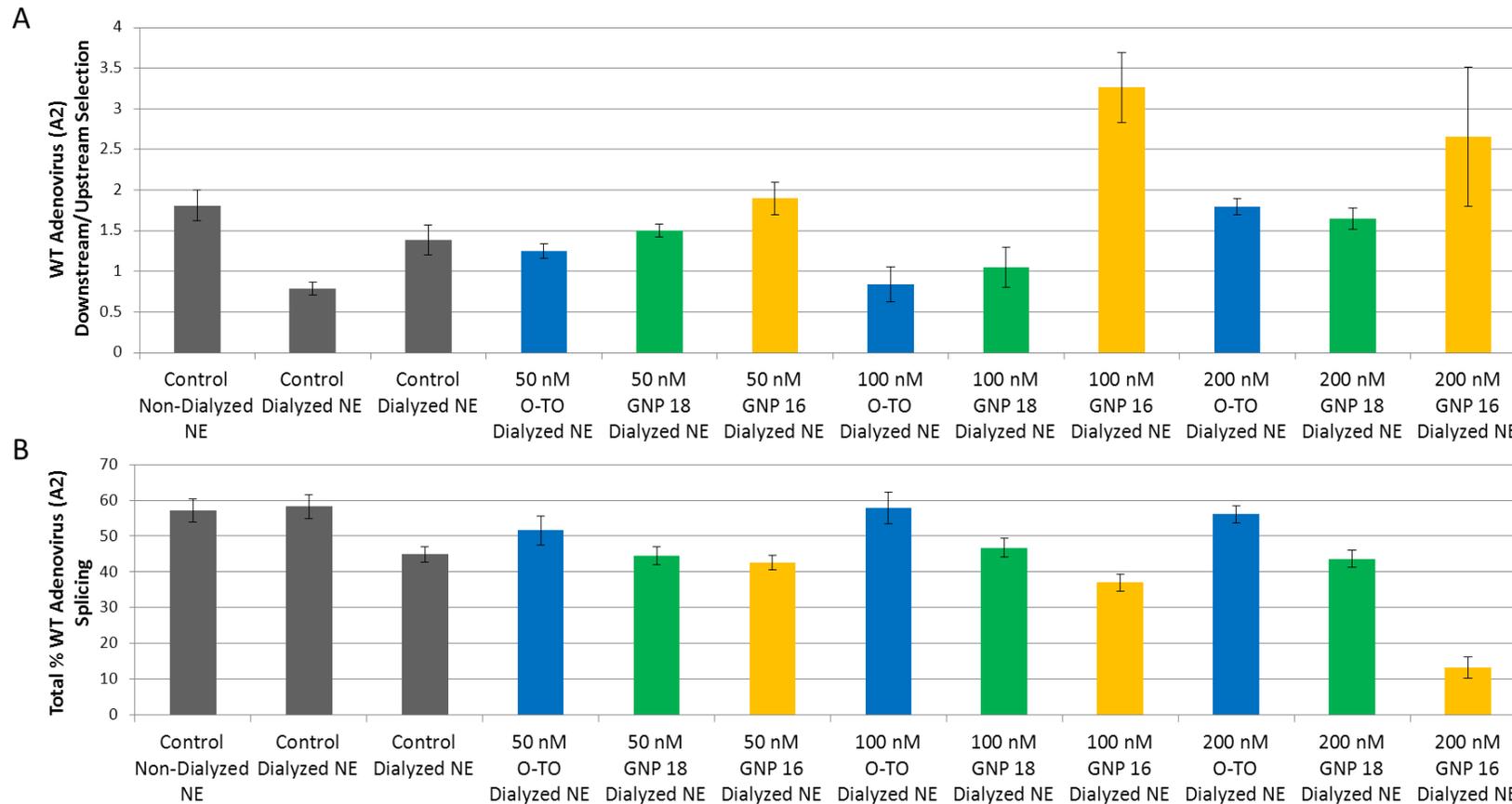


Figure 5.5: Quantification of splicing with GNP 16 using the data taken from the experiment in Figure 5.4

(A) Graph showing the ratio of mRNA products resulting from downstream versus upstream 5' splice site selection. GNP 16 clearly promoted use of the downstream splice site, and this effect increased as the concentration of GNP 16 was increased. **(B)** Graph showing the total percentage of A2 splicing, determined using the pre-mRNA and mRNA product values.

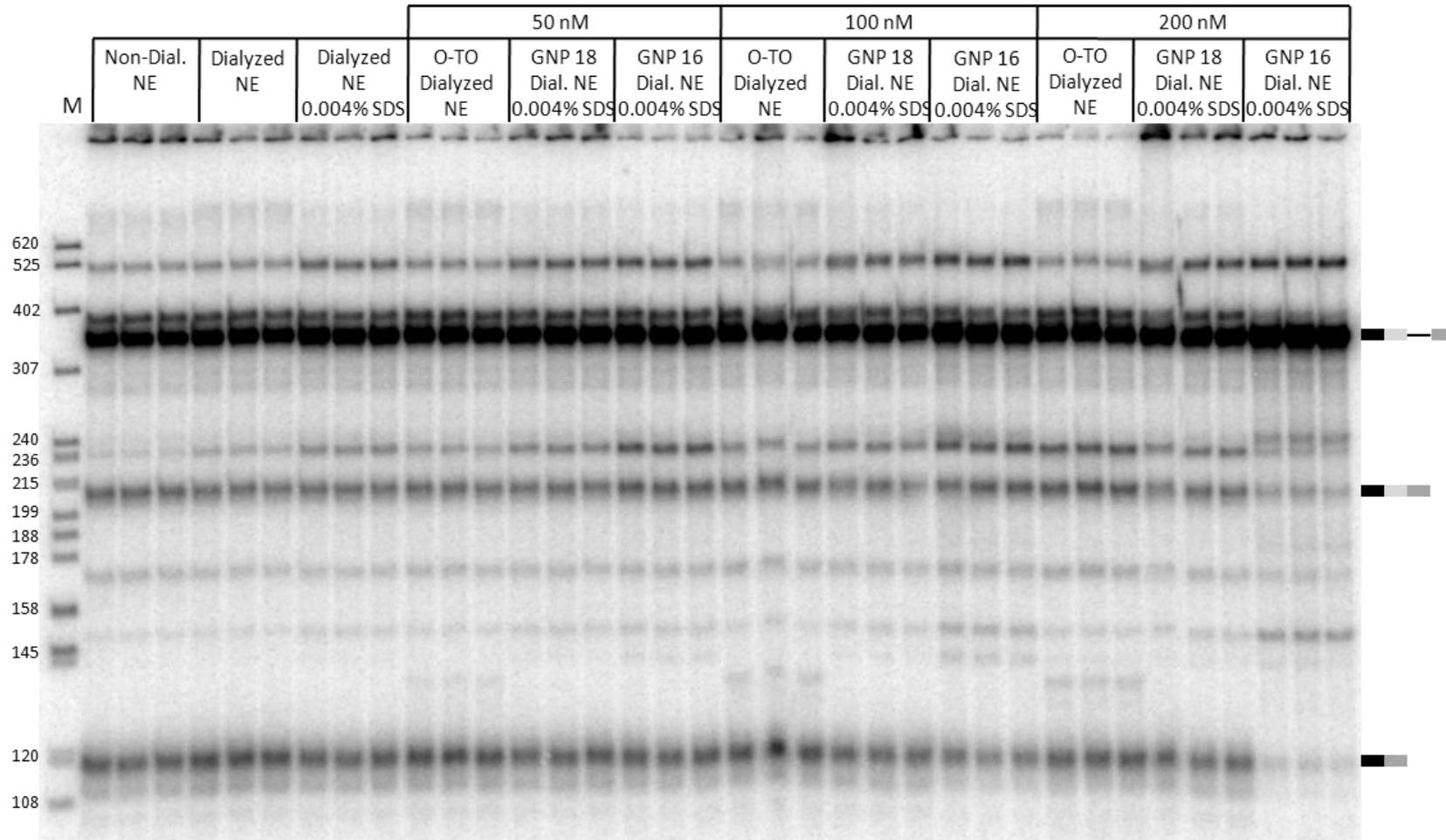


Figure 5.6: Denaturing PAGE of triplicate *in vitro* splicing of adenovirus-derived transcript A3 in HeLa nuclear extract at 2 hrs with increasing concentrations of GNP 16. Quantification (See Section 2.2.7) and analysis of these results can be seen in Figure 5.7.

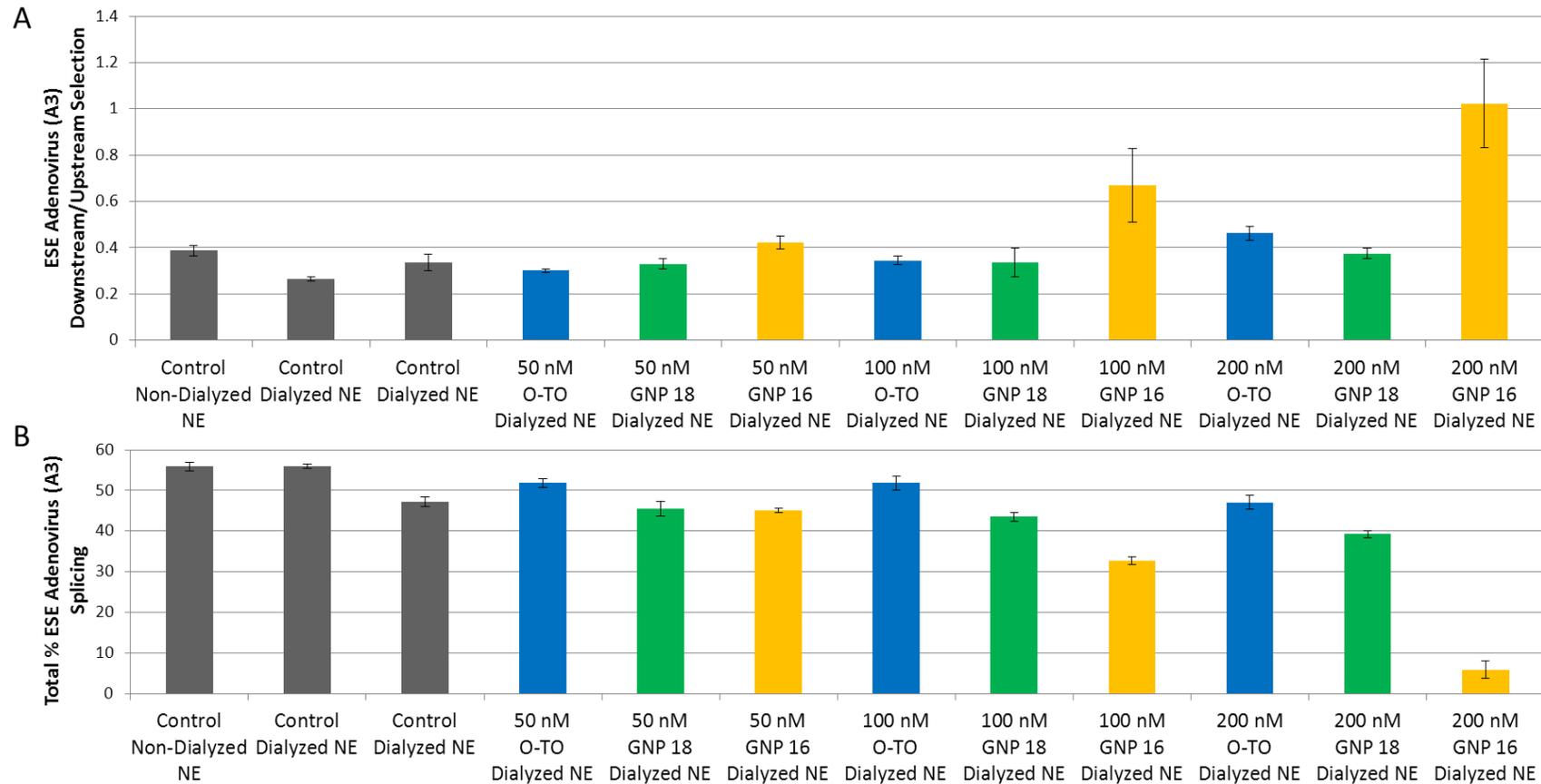


Figure 5.7: Quantification of splicing with GNP 16 using data taken from the experiment in Figure 5.6.

(A) Graph showing the ratio of mRNA products resulting from downstream versus upstream 5' splice site selection. GNP 16 clearly promoted use of the downstream splice site, and this effect increased as the concentration of GNP 16 was increased. **(B)** Graph showing the total percentage of A3 splicing, determined using the pre-mRNA and mRNA product values.

5.3 *In vitro* Splicing with Morpholino Oligonucleotides

Targeting a Silencer in SMN2 Intron 7

Several regulatory elements, both positive and negative, have been found to modulate SMN2 splicing (Section 1.3.1). In response to this, several different oligonucleotide based strategies to rescue the splicing of SMN2 exon 7 have been developed, which target some of these regulatory elements (Section 1.3.2). One such strategy, which has been found to be successful, involves blocking the splicing silencer N1 in SMN2 intron 7 (ISS-N1) with various antisense oligonucleotides (Hua *et al.*, 2011; Rigo *et al.*, 2012; Porensky *et al.*, 2012; Zhou *et al.*, 2013). Morpholino oligonucleotides have been found in both animal models and clinical trials to be both effective and well tolerated with no adverse effects being reported (Sazani *et al.*, 2011; Cirak *et al.*, 2011; Porensky *et al.*, 2012). To test whether such oligonucleotides might be effective in SMA, three oligonucleotides of various lengths were tested for their ability to suppress ISS-N1. This section looks at the results of the *in vitro* SMN2 splicing with these three morpholino oligos (PMO18, PMO20, and PMO25). Additional assays in SMA patient fibroblasts and a severe mouse model of SMA were done by collaborators, and the results have been published (Zhou *et al.*, 2013).

In vitro SMN2 splicing with PMO18, PMO20 and PMO25 at concentrations of 50 nM, 100 nM, 250 nM and 500 nM, done in triplicate, showed that all three oligos were capable of concentration dependent stimulation of SMN2 exon 7 inclusion (Figure 5.8). Interestingly, all three showed a dramatic increase in efficacy at 250 nM. Whilst all three oligos were clearly functional, quantification revealed that PMO25 was superior

to PMO18 and PMO20 (Figure 5.9). At 250 nM, PMO25 showed a slightly greater total level of SMN2 exon 7 inclusion (Figure 5.9A) and an approximate 1.8 fold increase in the ratio of exon 7 inclusion/exclusion, compared to PMO18 and PMO20 (Figure 5.9B). These *in vitro* results were consistent with the *in vivo* results (Zhou *et al.*, 2013).

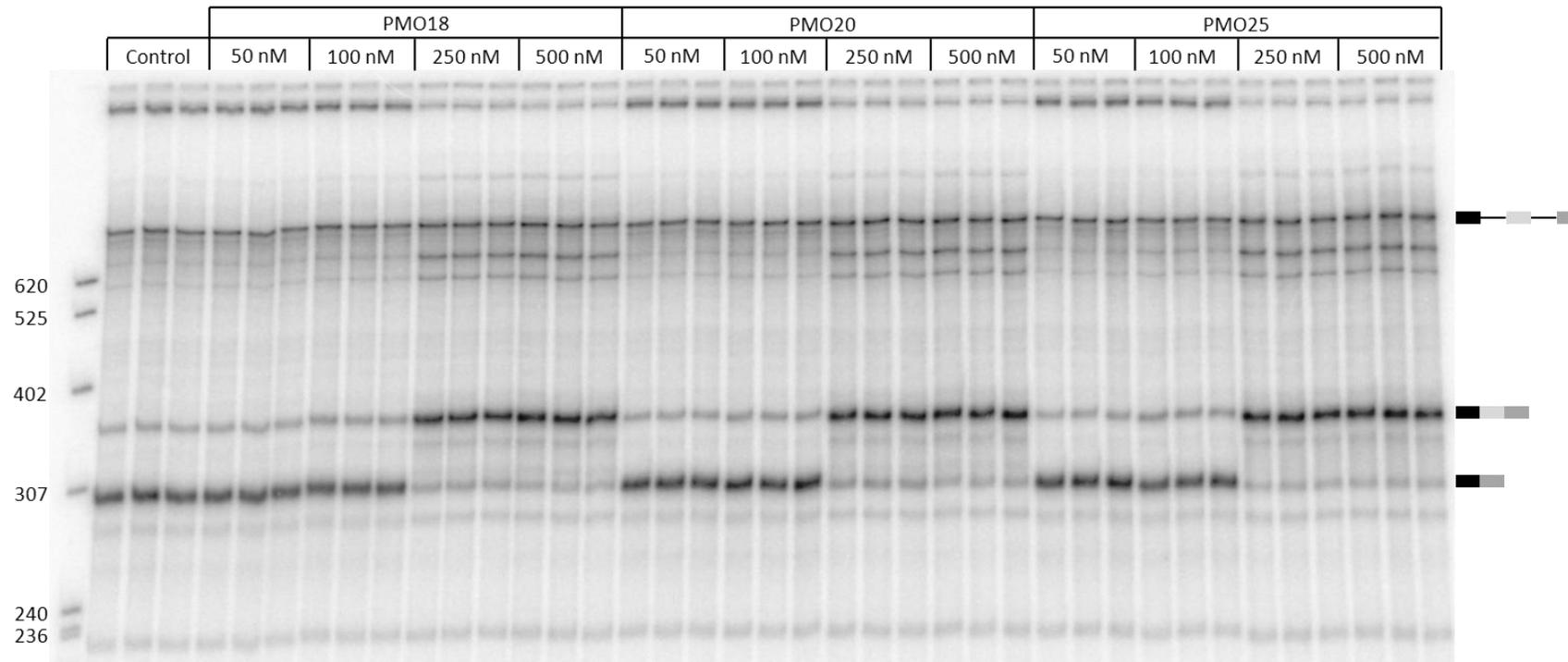


Figure 5.8: Denaturing PAGE of triplicate *in vitro* SMN2 splicing in HeLa nuclear extract at 2 hrs with increasing concentrations of three antisense morpholino oligonucleotides targeting an ISS in exon 7 (PMO18, PMO20, PMO25). Quantification (See Section 2.2.7) and analysis of these results can be seen in Figure 5.9.

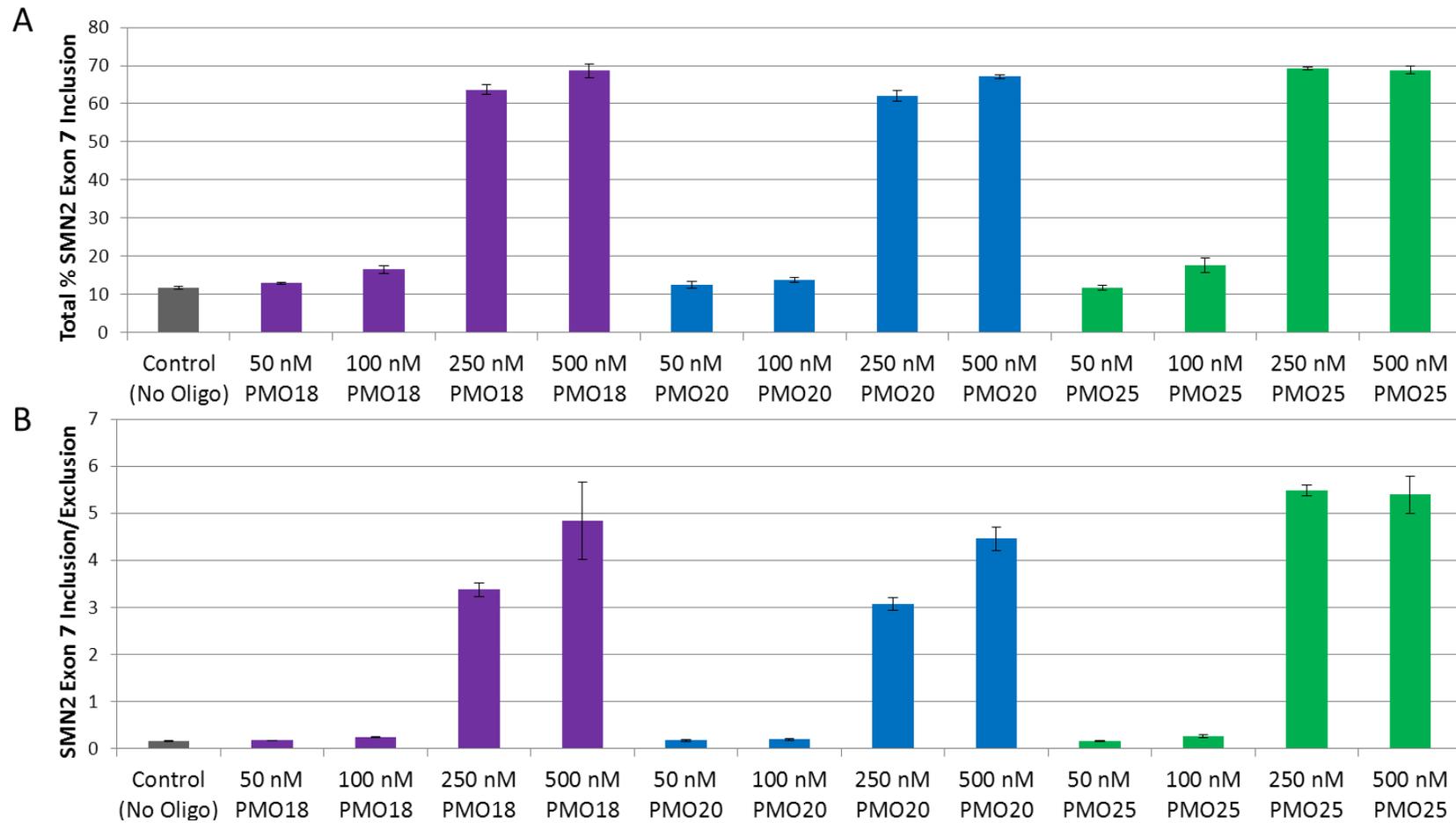


Figure 5.9: Quantification of the SMN2 *in vitro* splicing results with PMO18, PMO20, and PMO25 from Figure 5.8. PMO25 is superior at 250 nM. It gives a slightly higher total amount of exon 7 included mRNA and a reduced level of excluded product compared to the other oligos. (A) Graph showing the total percentage of SMN2 exon 7 inclusion. (B) Graph showing the ratio of SMN2 exon 7 inclusion/exclusion.

5.4 Discussion

New variations of two different antisense oligonucleotide (ASO) strategies for the stimulation of SMN2 exon 7 inclusion were investigated, using *in vitro* splicing. One of the new methods, based on the bifunctional targeted oligonucleotide enhancer of splicing (TOES) strategy developed in 2003 (Skordis *et al.*, 2003), involved the conjugation of gold nanoparticles (GNPs) with the annealing and enhancer domains of the TOES. The 5 nm gold nanoparticles were conjugated with over 400 strands of either a mix of the annealing and enhancer domains or just the enhancer domain, greatly increasing the localized concentration of oligonucleotide (Perrett *et al.*, 2013). Unexpectedly, the GNPs conjugated with only the enhancer domain (GNP 16) were more efficacious than those conjugated with both the enhancer and annealing domains (GNP 17). This indicated a non-specific effect because without the annealing domains the GNPs could not be specifically targeted to exon 7. Nevertheless, the stimulation of exon 7 splicing by GNP 16 was impressive, providing at least a proof of concept that oligonucleotide-conjugated GNPs can successfully alter splicing patterns. In fact, at lower concentrations (50 and 100 nM), GNP 16 even outperformed GGA-O, the relevant TOES control, at its best concentration of 250 nM. This non-specific enhancing effect was also observed when GNP 16 was tested with different, adenovirus-derived splicing substrates. Concentration seems to be a very important factor when using GNPs to modify splicing, as higher concentrations of GNPs were found to slow splicing significantly. Also, colleagues found that the size of the GNPs is important, with larger nanoparticles being more inhibitory (Perrett *et al.*, 2013). The GNPs used in this study were the smallest tested, measuring around 5 nm. Although a

non-specific effect on splicing is not ideal for therapeutic applications, it is an important finding. It proves that under the right circumstances, ASOs conjugated to gold nanoparticles can be successfully used to alter splicing patterns. This provides another interesting design possibility for future antisense oligonucleotide based therapeutic strategies.

The second new design variation tested involved blocking ASOs targeting the intronic splicing silencer (ISS) in intron 7 (ISSN-1). Other oligonucleotides targeting the same region have been tested previously and shown to be successful (Hua *et al.*, 2010; Hua *et al.*, 2011; Passini *et al.*, 2011; Porensky *et al.*, 2012). In fact, one of them has even shown success in phase 1 clinical trials in infants and has progressed to phase 2 trials recently (ClinicalTrials.gov Identifier NCT01494701, (News-medical.net, 2013)). The oligonucleotide in clinical trials, ISIS SMNRx, is an 18 mer 2'-O-(2-methoxyethyl) (MOE) phosphorothioate-modified ASO that blocks ISSN-1 (-10, -27) (Hua *et al.*, 2011). MOE phosphorothioate oligonucleotides, although widely used in clinical trials, have shown a potential for liver toxicity (Raal *et al.*, 2010; Kling, 2010; Akdim *et al.*, 2011).

However, a recent clinical trial of a phosphoramidate morpholino (PMO) ASO for the treatment of muscular dystrophy showed no such adverse effects (Cirak *et al.*, 2011).

In an effort to create an improved ASO, three PMO oligonucleotides of increasing length (PMO18, PMO20, PMO25), all annealing from the -10 position of intron 7, were created (Zhou *et al.*, 2013). *In vitro* SMN2 splicing assays comparing the three PMOs showed that increasing the length to 25 nt improved efficacy. At 250 nM concentration, PMO25 showed more total exon 7 inclusion and a reduced level of excluded product, giving a ratio of inclusion over exclusion that was clearly superior to

PMO20 and PMO18. These results were consistent with the *in vivo* data obtained by other colleagues (Zhou *et al.*, 2013), indicating a potential improvement in an already very promising ASO therapy for SMA.

6 USE OF Fe-BABE TO PROBE SPLICING REACTION MECHANISMS

6.1 Introduction

6.2 Synthesis of Fe-BABE Conjugated Oligonucleotides

6.3 Confirmation of Cleavage Conditions using Fe-EDTA

6.4 Fe-BABE Experiments on SMN2 using Fe-BABE Modified Oligonucleotides

6.5 Discussion

6.1 Introduction

The use of FeBABE as a tool for the investigation of splicing mechanisms was pioneered in 2002 (Kent & MacMillan, 2002). There is much potential for investigation of the mechanisms of alternative splicing regulation using FeBABE conjugated to various splicing components, including pre-mRNA, SR proteins, snRNPs, and repressor proteins. Oligonucleotides targeted to certain pre-mRNA areas, such as those being considered for therapeutic purposes, would also be good subjects for FeBABE studies. FeBABE is akin to the FeEDTA chelate used previously for footprinting assays, but has an extra group which allows it to be tethered to substrates containing a thiol group (Greiner *et al.*, 1997; Baichoo & Heyduk, 1997). It works in the same way as FeEDTA to catalyze the production of hydroxyl radicals in the presence of hydrogen peroxide and a reducing agent, such as ascorbic acid (Joseph & Noller, 2000). These hydroxyl radicals are highly reactive and are likely to have been consumed before they have diffused more than about 20 Ångströms (Dreyer & Dervan, 1985; Moser & Dervan, 1987; Kent & MacMillan, 2002). They will react with proteins, DNA and RNA resulting in cleavage. Because the hydroxyl radicals will react within such a short distance of their origin, tethered FeBABE studies can provide maps of close proximity interactions between the FeBABE tethered substrate and any number of targets (Joseph & Noller, 2000). The fact that the hydroxyl radicals will cleave any DNA or RNA with no sequence specificity makes FeEDTA a useful tool for mapping protection of DNA or RNA by the binding of proteins and possibly other strands of DNA or RNA. FeBABE and FeEDTA could be used in a number of ways to map out interactions of various splicing factors, enhancers,

silencers, and therapeutic oligonucleotides. In this study, Fe-BABE conjugated oligonucleotides were created to investigate the mechanism of action of GGA.

6.2 Synthesis of Fe-BABE Conjugated Oligonucleotides

In order to synthesize Fe-BABE modified oligonucleotides, it was first necessary to obtain oligonucleotides modified with a suitable reaction group for conjugation to the bromoacetamido group of Fe-BABE. To this end, 5' thiol modified oligonucleotides were synthesized by solid phase synthesis (Section 8.2.2). In order to conjugate Fe-BABE, first the di-thiol protection group had to be reduced with either DTT or TCEP, as shown in Figure 6.1A. TCEP, which was found to be effective at lower concentrations, was ultimately used in preference to DTT. Reduction conditions were confirmed by reverse phase HPLC analysis. TCEP was removed by dialysis immediately prior to conjugation with Fe-BABE, as it was found to react with the Fe-BABE. Then the deprotected 5' thiol-modified oligonucleotide was conjugated to Fe-BABE, as shown in Figure 6.1B. The Fe-BABE conjugation was followed by oligonucleotide mass changes resulting from deprotection and subsequent Fe-BABE modification, which were confirmed using high resolution polyacrylamide gel electrophoresis (Figure 6.1C).

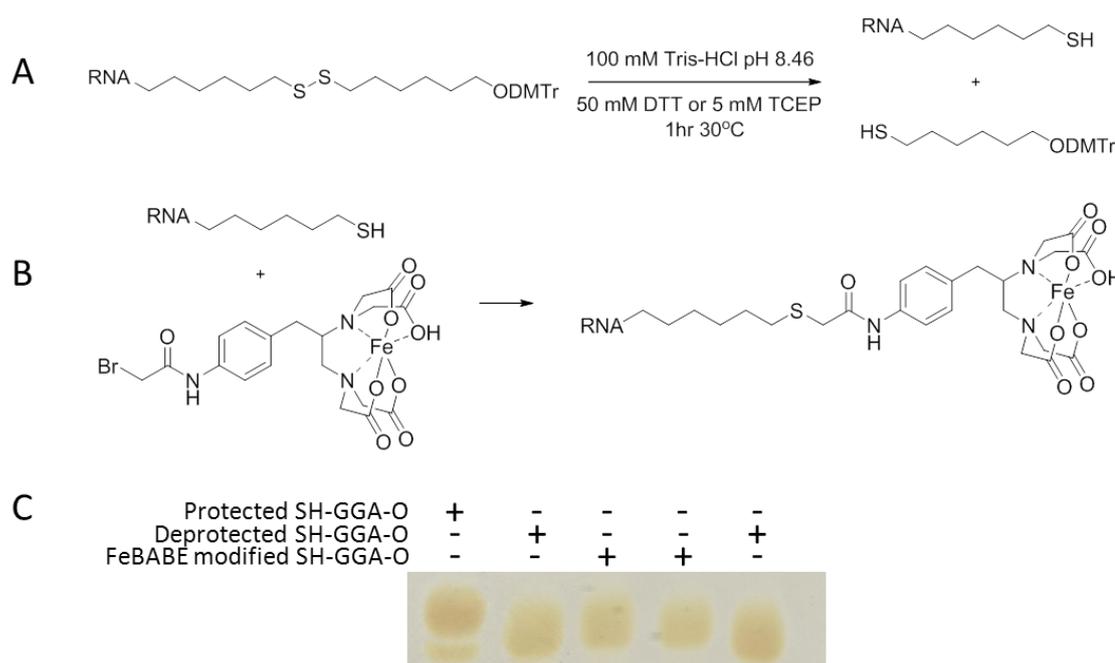


Figure 6.1: Conjugation of 5' thiol modified oligonucleotides with Fe-BABE.

(A) Schematic diagram illustrating the method for deprotection of the dithiol protecting group on the 5' end of the thiol modified oligonucleotides. **(B)** Schematic diagram showing the conjugation of Fe-BABE to the deprotected 5' thiol modified oligonucleotides. **(C)** Silver stained high resolution polyacrylamide gel showing mass changes of 5' thiol modified GGA-O (SH-GGA-O, Section 8.2.2) from deprotection of the dithiol protecting group and subsequent Fe-BABE conjugation.

6.3 Confirmation of Cleavage Conditions using Fe-EDTA

Prior to experimentation with Fe-BABE modified oligonucleotides, the conditions to trigger the production of hydroxyl radicals by Fe-BABE were confirmed using free Fe-EDTA, which has a similar structure and reacts in the same way as Fe-BABE. Many tests and optimizations were done using Fe-EDTA to cleave 5' radiolabelled GGA-O. The most important finding was the need to remove glycerol from all samples, as it has a very strong quenching effect on the radicals produced by Fe-EDTA (Figure 6.2). Dimethyl PEG was considered as a replacement for glycerol, as without the -OH groups, it would not have a quenching effect (Figure 6.2). However, the dimethyl PEG, when used in nuclear extract in place of glycerol, impaired the ability to assess

complex formation (data not shown). It was therefore decided to use nuclear extract dialyzed in glycerol-free buffer D, which was shown in Figure 6.3 to retain the ability to form a GGA dependent exon-defined complex on an SMN2 construct containing exon 7 with flanking intronic sequences (Section 8.1.2).

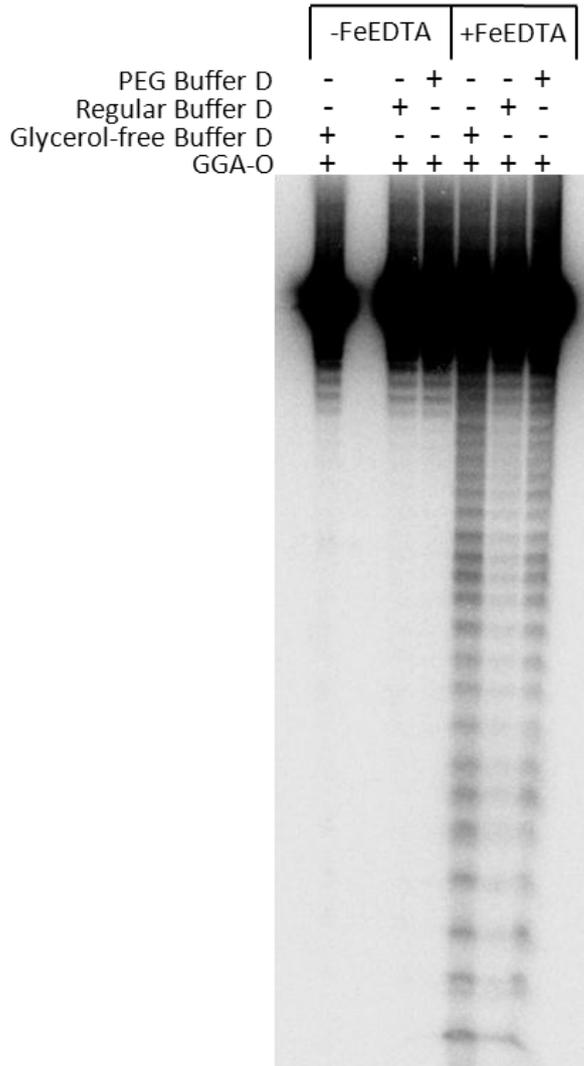


Figure 6.2: High resolution polyacrylamide gel showing cleavage of 5' radiolabelled GGA-O with free Fe-EDTA under different buffer conditions.

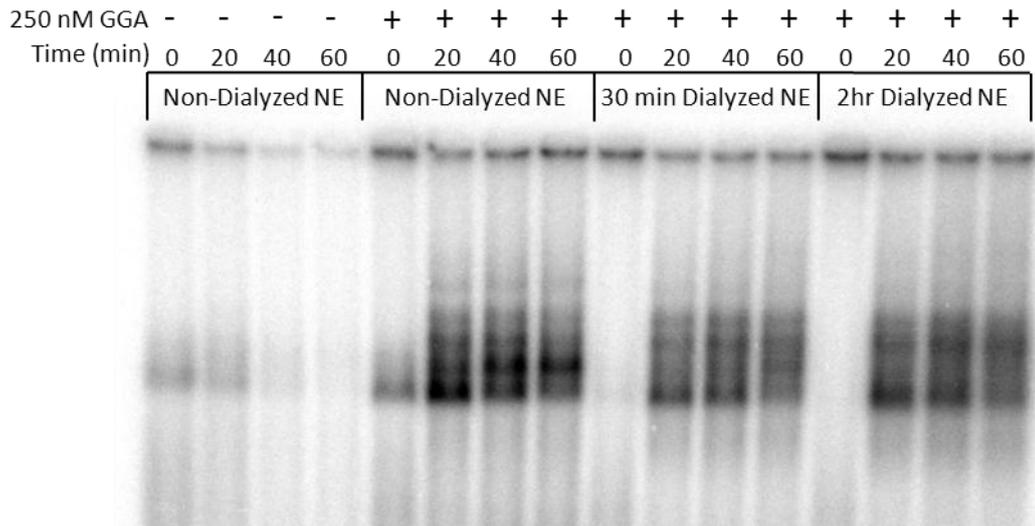


Figure 6.3: Native agarose complex-formation gel with radiolabelled SMN2 exon 7 defined construct (Section 8.1.2) in HeLa nuclear extract showing the effects of dialysis for either 30 min or 2 hrs in glycerol-free Buffer D on the ability to form a GGA dependent complex.

6.4 Fe-BABE Experiments on SMN2 using Fe-BABE

Modified Oligonucleotides

The first Fe-BABE experiments on SMN2 were done with Fe-BABE modified GGA-O (Fe-BABE GGA-O, Section 8.2.2) and an Fe-BABE modified 2'OMe control oligonucleotide containing the same annealing region as GGA-O and a tail of the same length with a sequence designed (using ESE finder (Cartegni *et al.*, 2003)) and confirmed by *in vitro* SMN2 splicing (data not shown) to have no enhancer properties (Fe-BABE NET-O, Section 8.2.2). The Fe-BABE modified oligonucleotides were annealed to cold SMN2 prior to incubation in nuclear extract under conditions causing stalling at splicing complex A. The production of hydroxyl radicals was initiated by addition of hydrogen peroxide and ascorbate, and then quenched using glycerol. The SMN2 RNA was purified and analyzed by reverse transcription with four primers to get maximal coverage and resolution of the sequence. However, no cleavage of SMN2 by these Fe-BABE modified oligonucleotides was observed in the absence of nuclear extract (Figure 6.4) or in the presence of nuclear extract dialyzed with glycerol-free Buffer D (Figure 6.5).

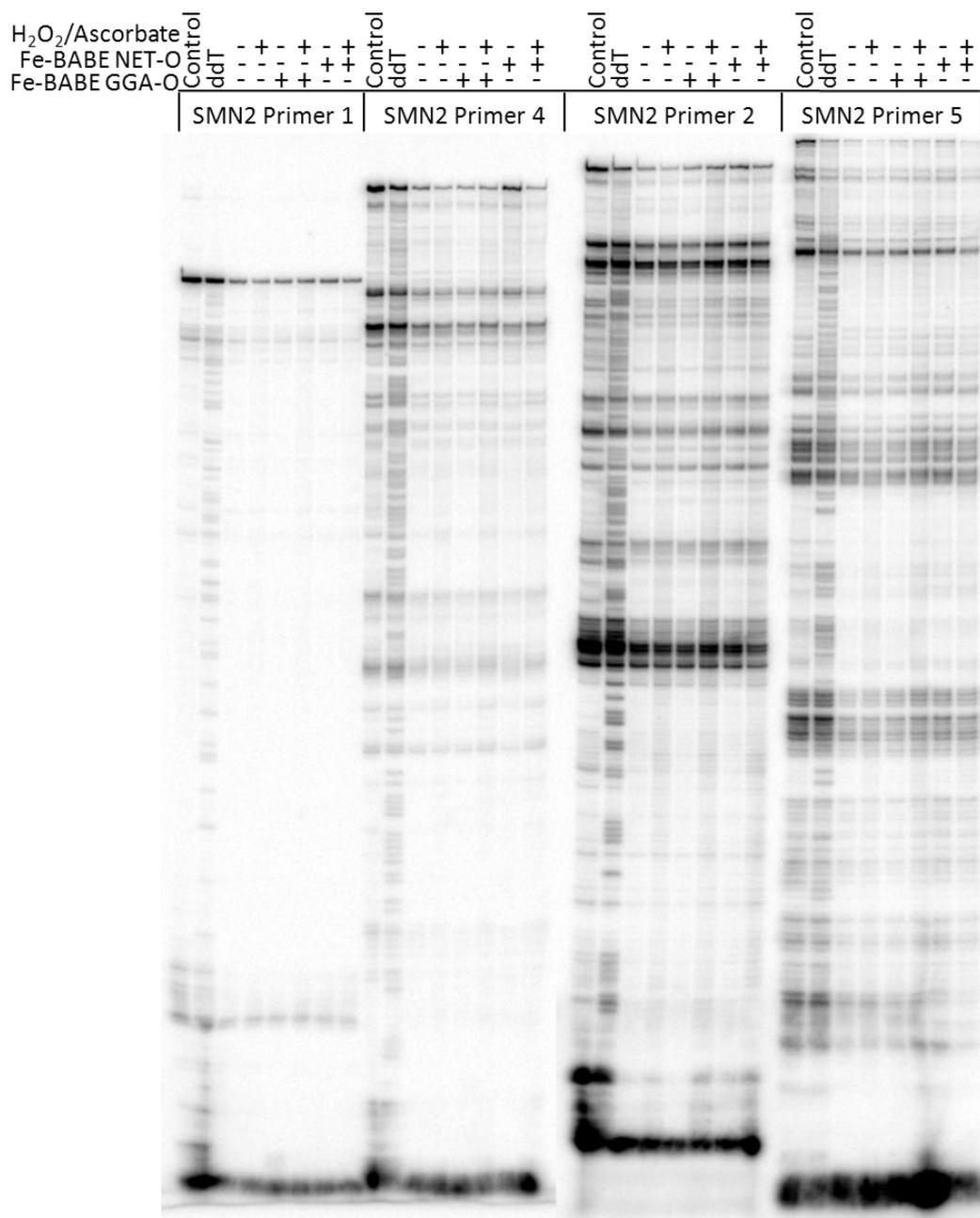


Figure 6.4: High resolution polyacrylamide gel showing reverse transcription of SMN2 RNA, which had been previously subjected to cleavage by Fe-BABE GGA-O and Fe-BABE NET-O, with four different 5' radiolabelled primers (Section 8.2.3). Fe-BABE experiment was done with typical splicing conditions, except that glycerol-free buffer D was used instead of nuclear extract. No specific Fe-BABE cleavage was observed for any samples.

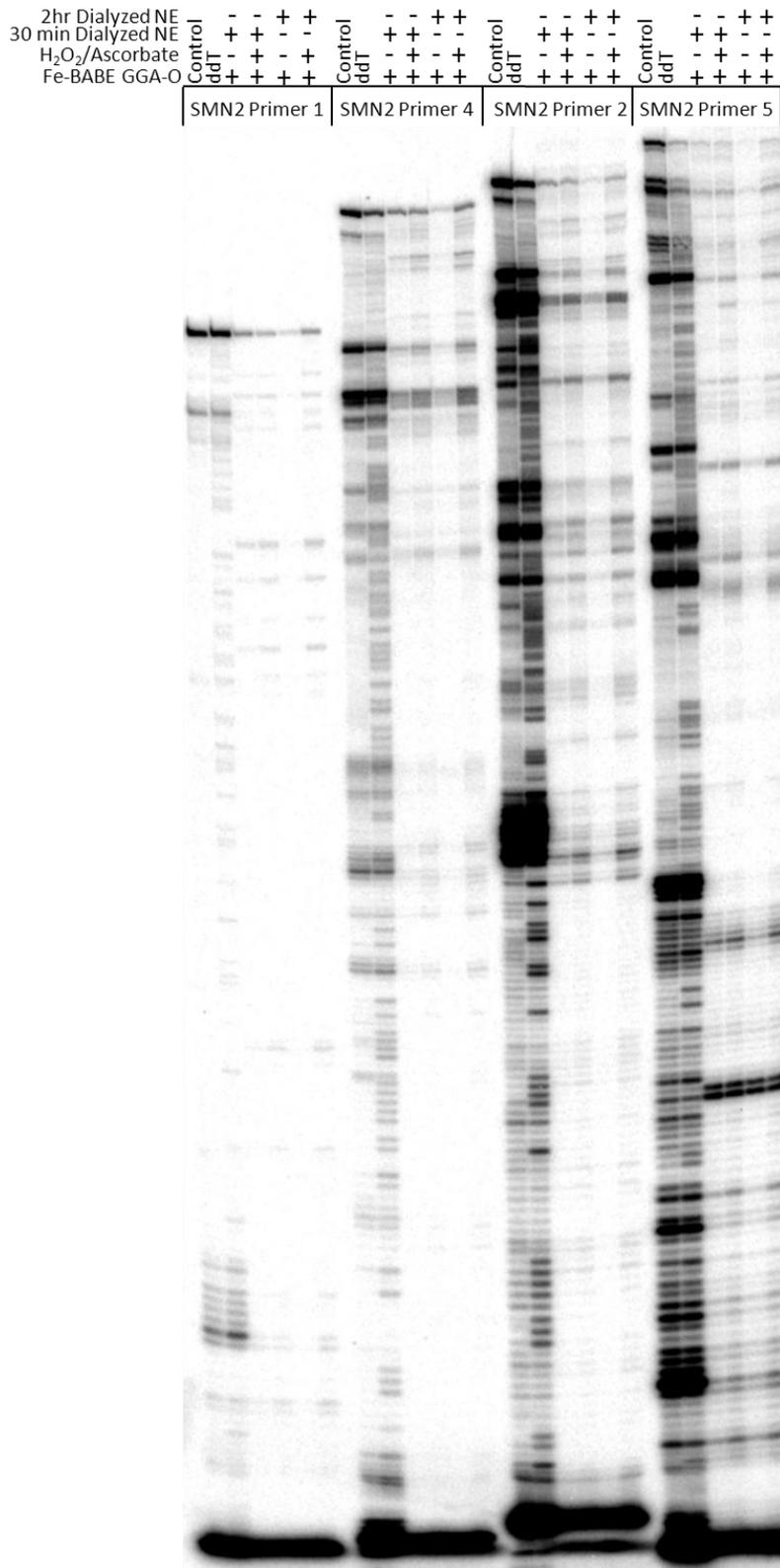


Figure 6.5 High resolution polyacrylamide gel showing reverse transcription of SMN2 RNA from Fe-BABE experiment using four different 5' radiolabelled primers (Section 8.2.3). Fe-BABE GGA-O experiment was done under A complex conditions with HeLa nuclear extract dialyzed in glycerol-free Buffer D. No specific cleavage was observed.

As Fe-BABE GGA-O and Fe-BABE NET-O were unable cause cleavage of SMN2 RNA under the conditions determined by Fe-EDTA experiments, further optimization of the cleavage conditions was done with radiolabelled SMN2 and Fe-BABE GGA-O (Figure 6.6). However, even with increased levels of hydrogen peroxide and ascorbate, no specific cleavage of SMN2 in the presence of Fe-BABE GGA-O was observed. At this point it was hypothesized that the Fe-BABE might be too far away from the SMN2 transcript to cause significant cleavage, because it was conjugated at the end of the 23 nt tails of Fe-BABE GGA-O and Fe-BABE NET-O. Also, the fully 2'OMe GGA-O was found to be less effective at stimulating SMN2 exon 7 inclusion. Therefore, new Fe-BABE modified oligonucleotides, Fe-BABE NT and Fe-BABE GGA (Section 8.2.2) were synthesized with the same chemical modifications as GGA. Fe-BABE NT, which as it has no enhancer tail would position the Fe-BABE very near to the SMN2 transcript, was made to test the hypothesis described previously. As a preliminary experiment, radiolabelled SMN2 was subjected to cleavage by a large excess of Fe-BABE NT and Fe-BABE GGA (Figure 6.7). Cleavage was only observed in the presence of Fe-BABE NT under cleavage conditions, indicating that the position of Fe-BABE at the end of a 23 nt tail may have been to blame for the lack of cleavage observed in previous experiments. It was strange that a smear of cleaved SMN2 RNA was observed rather than two, more distinct, bands. However, as no cleavage was observed with Fe-BABE GGA, even in such a large excess, further experimentation seemed futile.

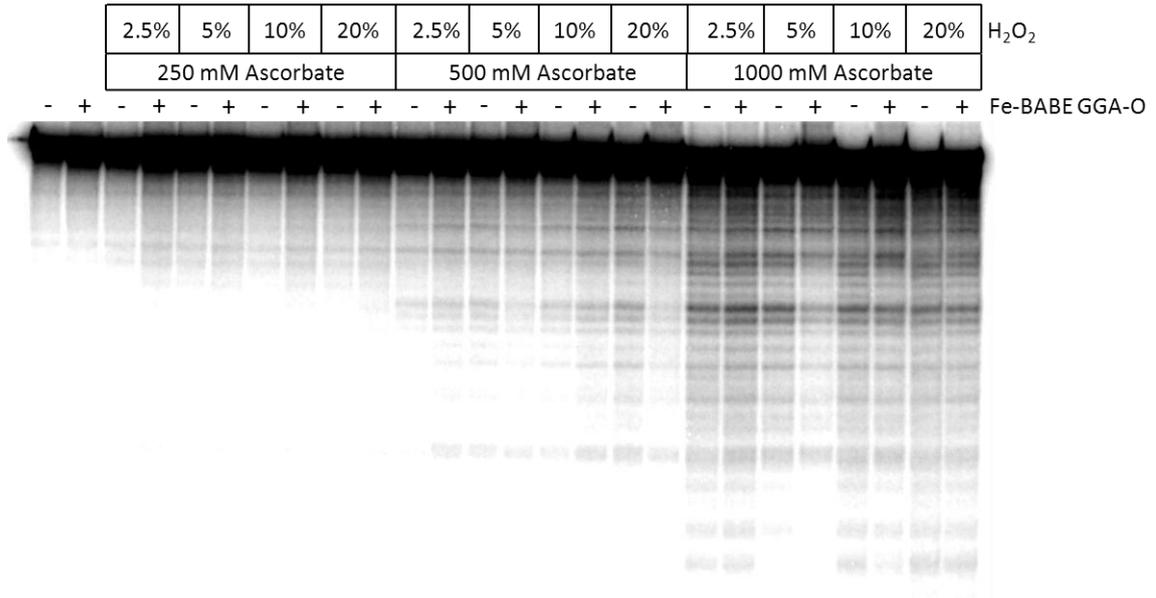


Figure 6.6: Optimization of cleavage conditions with Fe-BABE GGA-O and radiolabelled SMN2 RNA

As shown by denaturing PAGE, despite increasing the concentrations of ascorbic acid and hydrogen peroxide to very high levels, only non-specific cleavage was observed. It is important to note that the concentrations of hydrogen peroxide and ascorbate shown are not final concentrations, but rather the concentrations of the stocks added to initiate the reaction.

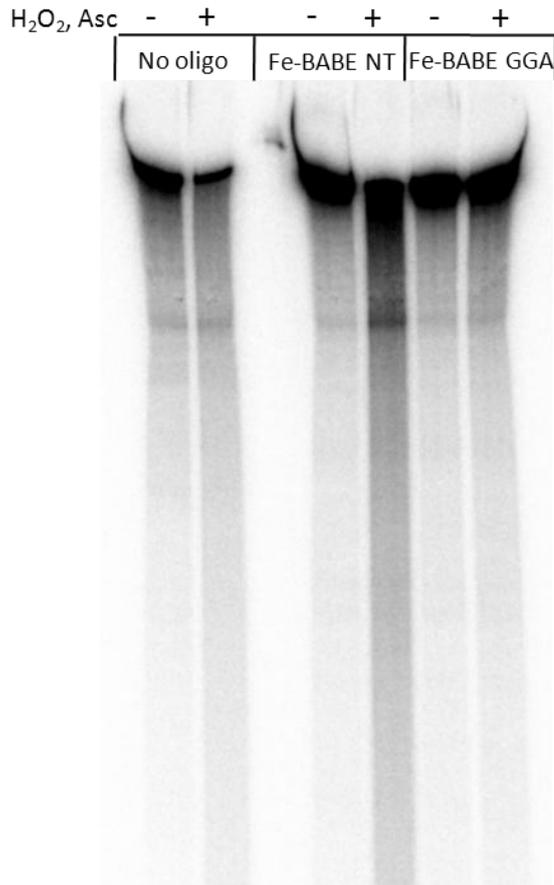


Figure 6.7: Radiolabelled SMN2 cleavage test using Fe-BABE NT and Fe-BABE GGA

Cleavage was only observed by denaturing PAGE with Fe-BABE NT, which positions the Fe-BABE group right next to the SMN2 RNA. Fe-BABE GGA, which has the Fe-BABE group at the end of its 23 nt enhancer tail, did not show any cleavage.

6.5 Discussion

In order to investigate the mechanism of action of GGA, a bifunctional oligonucleotide enhancer of SMN2 exon 7 splicing, Fe-BABE modified oligonucleotides were synthesized. FeBABE is an iron chelating moiety that, when activated with hydrogen peroxide and ascorbic acid, generates highly reactive hydroxyl radicals capable of cleaving DNA, RNA and proteins within a very short radius of their origin. FeBABE modified versions of GGA-O, the 2'OMe version of GGA, and a control oligo with a non-enhancer tail of the same length as GGA-O (NET-O) were synthesized (See Section 8.2.2 for oligo details.). Despite extensive optimizations, no Fe-BABE induced cleavage was ever observed using these oligonucleotides. It was hypothesized that the positioning of the FeBABE group at the end of the 23 nt tails could be at fault by not allowing the hydroxyl radicals to be in close enough proximity to the target SMN2 RNA to achieve a detectable level of cleavage. To test this, two new FeBABE modified oligonucleotides were created: Fe-BABE GGA, containing the same sequence as GGA-O but more efficacious chemical modifications, and Fe-BABE NT, an oligonucleotide containing only the 15 nt annealing region and lacking the 23 nt tail (See Section 8.2.2 for oligo details.). A cleavage test with Fe-BABE NT and Fe-BABE GGA showed that only Fe-BABE NT, which positions the Fe-BABE group directly next to the SMN2 transcript, produced any cleavage. Unfortunately, this indicated that further experimentation with Fe-BABE GGA was unlikely to succeed and different methods needed to be used to investigate the mechanism of action of GGA.

7 DISCUSSION

7.1 Importance of Understanding the Mechanism by which GGA Enhances SMN2 Exon 7 Inclusion

7.2 Stabilization of the G-quadruplex Structure of the GGA Enhancer Tail Domain Decreases Functionality

7.3 GGA Directly Binds SRSF1 and G-quadruplex Associated Proteins and Associates with many Splicing and Transcription-Related Proteins

7.4 The Enhancer Tail and Annealing Domains of GGA both have Significant Mechanistic Roles

7.5 Overview of Proposed Mechanism of Action of GGA

7.1 Importance of Understanding the Mechanism by which GGA Enhances SMN2 Exon 7 Inclusion

Splicing is a highly complex, highly regulated cellular process that is very susceptible to alteration via mutations. In fact, up to 50% of disease-causing mutations are thought to affect splicing (López-Bigas *et al.*, 2005). The use of antisense oligonucleotides to alter splicing patterns is a very promising therapeutic strategy for diseases resulting from changes to splicing. These oligonucleotides are designed to either block or recruit splicing regulatory factors, thereby influencing splice site usage. However, determining the precise mechanism by which these oligonucleotides influence splice site selection is non-trivial, as the natural regulatory mechanisms exploited by the oligonucleotides are still not fully understood.

GGA is a targeted oligonucleotide enhancer of splicing (TOES), which was designed in 2003 as a potential therapeutic agent for spinal muscular atrophy (Skordis *et al.*, 2003). It is a bifunctional oligonucleotide, which has an annealing domain that targets it to SMN2 exon 7 and a non-annealed enhancer tail containing three repeats of GGAGGAC. It was designed to stimulate exon 7 inclusion by recruiting activator proteins, such as SRSF1, with this enhancer domain. Initial results *in vitro* and in SMA patient fibroblasts showed that GGA was indeed successful at stimulating SMN2 exon 7 inclusion (Skordis *et al.*, 2003). Additionally, when this same sequence was incorporated into a U7 snRNA and transgenically expressed in a mouse model of spinal muscular atrophy, the mice were essentially able to live normal lives (Meyer *et al.*, 2009). Clearly, this and similar oligonucleotides have great potential therapeutically. However, when attempts were

DISCUSSION

made to improve the design of this oligonucleotide, they were only successful in providing a guide to future oligonucleotide designs, not in improving the original design (Owen *et al.*, 2011). This indicates that increased knowledge of the precise mechanism of action of this oligonucleotide is necessary in order to more easily and efficiently design bifunctional TOES for other diseases. Determination of the mechanism of action of GGA would also provide further insights into how enhancer sequences in general work.

7.2 Stabilization of the G-quadruplex Structure of the GGA Enhancer Tail Domain Decreases Functionality

NMR and Circular Dichroism (CD) studies done by colleagues from the University of Nottingham (Mr A. Cousins and Prof. M. Searle) with the help of Mr A. Perrett (University of Leicester) showed that the enhancer tail domain of GGA (AGGAGGACGGAGGACGGAGGACA) forms a G-quadruplex (Smith *et al.*, manuscript submitted). In fact, it is likely that, at least some of the time, a stacked dimer of GGA tail quadruplexes is formed (Smith *et al.*, manuscript submitted). These findings sparked investigation into whether or not the G-quadruplex structure was present in a cellular context. Reverse transcription along the enhancer tail of GGA, both in buffer and in nuclear extract, gave a distinctive termination pattern, which was consistent with the presence a G-quadruplex (Figure 4.24A). Ideally, this would have been confirmed further by using a version of GGA containing 7-deaza-G nucleotides, which are unable to form the Hoogsteen base pairs necessary for G-quadruplex formation (Murchie & Lilley, 1994; Seela *et al.*, 1998). However, attempts to transcribe GGA using

DISCUSSION

7-deaza-rGTP were unsuccessful. After determining that the G-quadruplex structure was likely to form under biologically relevant conditions, interest was turned to the functional relevance of the G-quadruplex structure.

To investigate the effect of the G-quadruplex structure on the function of GGA, aromatic ligands designed to stabilize G-quadruplex structures were obtained from Professor Laurence Hurley (University of Arizona) (Figure 4.15). *In vitro* SMN2 splicing assays with these ligands and GGA showed a concentration dependent decrease in SMN2 exon 7 inclusion (Figure 4.17, Figure 4.18 and Figure 4.19). This effect was also seen to some degree in the control samples lacking GGA. However, for two of the ligands, GSA-0820 and GSA-0902, the reduction in exon 7 inclusion was proportionally much greater in the samples containing GGA, indicating that the G-quadruplex structure needs to melt prior to the binding of the activator protein(s) to the enhancer tail domain (Figure 4.20). Interestingly, UV crosslinking and native complex gel analysis of full length GGA and the enhancer tail domain (TO) showed that protein binding and complex formation was not significantly altered in the presence of the quadruplex stabilizing ligands (Figure 4.21, Figure 4.22 and Figure 4.23). This most likely indicates that the enhancer tail domain exists predominantly in the G-quadruplex form when not in an active splicing complex. This could be a useful property, which could even be used in future TOES designs, as it might prevent sequestration of activator proteins by the free oligonucleotide when administered therapeutically.

7.3 GGA Binds SRSF1 and G-quadruplex-Associated Proteins and Recruits many Splicing and Transcription-Related Proteins

Identification of the proteins that bind to and associate with GGA was a vital step towards determination of its mechanism of action. When designed in 2003, the enhancer tail domain of GGA was designed with the consensus binding site for SRSF1 in mind (Skordis *et al.*, 2003). Therefore, it was logical to investigate SRSF1 binding. This was done in several ways, most conclusive of which was short wave UV crosslinking of 5' radiolabelled GGA in nuclear extracts made from cells expressing eGFP-tagged SRSF1 (Figure 4.5). Interestingly, crosslinking in eGFP-SRSF1 nuclear extract made from HEK 293T cells, which are known to highly express SR proteins like SRSF1, showed almost exclusive binding of eGFP-SRSF1 compared to the equivalent experiment done in eGFP-SRSF1 HeLa nuclear extract, which showed binding of several other proteins in addition to SRSF1 (Figure 4.5A and E). Binding of purified SRSF1 to 5' radiolabelled GGA was also detected using native polyacrylamide complex gels, but this interaction was very sensitive to experimental conditions (Figure 4.8 and Figure 4.9). These experiments provide clear evidence that SRSF1 does bind the enhancer tail domain of GGA. However, it was necessary to take an unbiased approach and investigate all of the proteins associated with GGA.

In order to get a complete view of all proteins associated with the enhancer tail of GGA, an affinity purification procedure was developed using a biotinylated version of GGA containing a photocleavable linker between the annealing domain and the

DISCUSSION

enhancer tail, GGA PC Bio (Chapter 3). This allowed for gentle purification and elution of the proteins associated with the enhancer tail domain of GGA. Analysis of a single SDS PAGE gel band containing all of the purified proteins was done using the sensitive method of liquid chromatography - tandem mass spectrometry (LC MS/MS), revealing 42 proteins unique to the +GGA PC Bio sample (Figure 4.1). The proteins identified fall into several categories including splicing regulatory proteins, spliceosomal proteins, transcriptionally-related factors, RNA helicases, and G-quadruplex associated proteins. Given the range and types of proteins found, it seems likely that the enhancer domain of GGA is capable of associating, via protein-protein interactions, to pre-mRNAs present in the nuclear extract. It is important to note that SRSF1 was the highest ranking activator protein (9th ranked overall for percent coverage) identified by the LC MS/MS analysis, giving further support to the expectation that it is associated with the enhancer domain and is likely to be responsible for the stimulation of SMN2 exon 7 inclusion by GGA.

Many proteins in the list have been shown to be in some way associated with sequences of DNA or RNA that can form G-quadruplexes. The highest ranking protein by percent coverage was cellular nucleic acid binding chaperone protein CNBP, which has been shown to promote formation of parallel G-quadruplexes (Michelotti *et al.*, 1995; Liu *et al.*, 1998; Armas *et al.*, 2008). HnRNP A1, ranked 3rd, is a splicing repressor protein also associated with telomeres (Zhang *et al.*, 2006). Ranked 4th is DHX36, an RNA helicase shown to promote unwinding of G-quadruplexes (Vaughn *et al.*, 2005; Creacy *et al.*, 2008; Giri *et al.*, 2011). XRN2, a nuclear 5'-3' exoribonuclease has not directly been shown to be associated with G-quadruplexes; however, its close relative

DISCUSSION

XRN1, a cytoplasmic 5'-3' exoribonuclease, which ranked 37th in the affinity purification, has been shown to have a preference for G-quadruplex forming substrates (Bashkirov *et al.*, 1997). Ranked 6th and 8th are two version of hnRNP H, a splicing regulatory protein shown to associate with G-triplets (Caputi & Zahler, 2001). Closely related to hnRNP H is hnRNP F (ranked 19th), which also binds G-triplets and has been found to promote unfolding of G-quadruplex forming sequences (Dominguez & Allain, 2006; Dominguez *et al.*, 2010). Nucleolin, ranked 16th, has been shown to stabilize parallel G-quadruplexes through its binding (Bates *et al.*, 1999; González *et al.*, 2009; González & Hurley, 2010; Indig *et al.*, 2012). Along with nucleolin, two other G-quadruplex-associated transcription factors, the transcriptional activator protein Pur-alpha (PURA) and interleukin enhancer-binding factor 2 (ILF2), ranked 17th and 20th respectively, were identified. All three of these proteins have recently been shown to be associated with the G-quadruplex forming GGA-rich sequence of the ERBB2 gene promoter (Zhang *et al.*, 2012). The abundance of G-quadruplex associated proteins in the affinity purification sample is rather interesting and fits with the finding that GGA exists largely in quadruplex form.

The LC MS/MS results clearly indicate that GGA is capable of associating with many different proteins, but many of these proteins are not binding directly to GGA. UV crosslinking to 5' radiolabelled GGA in HeLa nuclear extract shows two strong bands in addition to the weaker SRSF1 sized band (Figure 4.2B). Identification of these proteins was necessary in order to determine if other proteins besides SRSF1 could be responsible for the stimulatory effect GGA has on SMN2 exon 7 splicing.

DISCUSSION

The smaller band, which runs at about 30 kDa when crosslinked to GGA and 25 kDa when crosslinked to TO is consistent with the size of CNBP, when the additional shift from the crosslinked oligonucleotides is taken into account. UV crosslinking in HEK 293T nuclear extract with overexpressed, T7-tagged CNBP showed an increased intensity in this band for GGA and a small shift in addition to the increased intensity for TO (Figure 4.4C). The larger size of the GGA oligonucleotide made it impossible to detect a shift from the small T7 tag. However, the increased intensity in the band is consistent with the overexpression of T7-CNBP. These crosslinking results, in addition to the fact that CNBP was the highest ranking protein identified by LC MS/MS after affinity purification, indicate that CNBP binds directly to the enhancer tail domain of GGA. There are currently no results in the literature that would suggest CNBP could be responsible for the stimulation of splicing achieved by GGA.

The other major protein found to crosslink to GGA in addition to CNBP and SRSF1 was found to bind much more to the full length GGA compared to the enhancer tail domain on its own (Figure 4.2B). The identity of this protein was not confirmed, but it is likely to be nucleolin. A large scale affinity purification with GGA PC Bio (Chapter 3) followed by individual-band analysis using MALDI-TOF mass spectrometry identified the most abundant band to be nucleolin (Figure 4.3). The fact that this band is found to crosslink much less to the enhancer tail domain on its own when compared to full length GGA indicates that the annealing region of GGA may help to stabilize binding of this protein. This mode of binding is logical for nucleolin. As mentioned previously, nucleolin is known to bind G-quadruplexes. Another well-known G-quadruplex binding protein is fragile X mental retardation protein (FMRP). A peptide of the RGG domain of FMRP has

DISCUSSION

been shown by NMR to bind at the junction between an RNA G-quadruplex and connected duplex structure (Phan *et al.*, 2011). The RGG domain of nucleolin likely binds in a similar fashion. This would explain why it is found to crosslink less to the enhancer tail domain on its own. This is also consistent with fact that nucleolin was so abundant in the affinity purification. Nucleolin would have been able to bind stably to the full length GGA PC Bio, and when the enhancer tail domain was released by photocleavage, this would have destabilized the interaction at the junction of the quadruplex-forming enhancer domain and the annealing domain. Like CNBP, nucleolin has not been shown thus far to have a role as a splicing activator protein. Stabilization of the G-quadruplex structure of GGA with ligands led to a decrease in functionality (Section 7.2), further indicating that G-quadruplex binding proteins are not likely to be the 'active' proteins. Immunoprecipitation of the enhancer domain of GGA after UV crosslinking using an antibody against SRSF1 showed that nucleolin does not bind when SRSF1 is binding and that hnRNP F/H and CNBP bind at a reduced level (Mr C. Lucas personal communication, (Smith *et al.*, manuscript submitted)). A similar experiment with an antibody against hnRNP F/H showed that hnRNP F/H tends to bind exclusively, indicating that there are different protein complexes capable of forming on the enhancer tail of GGA (Mr C. Lucas personal communication, (Smith *et al.*, manuscript submitted)). These findings are consistent with the idea that the G-quadruplex form of GGA is not the 'active' form of GGA. Therefore, SRSF1, a well-known splicing activator protein, is most likely responsible for the stimulatory function of GGA on SMN2 exon 7 inclusion.

7.4 The Enhancer Tail and Annealing Domains of GGA both have Significant Mechanistic Roles

GGA is a bifunctional oligonucleotide that anneals from nucleotides 2-16 at the 5' end of SMN2 exon 7 and has an enhancer tail (AGGAGGACGGAGGACGGAGGACA) which extends up from position 16 (Skordis *et al.*, 2003). One of the key differences between SMN1 and SMN2, which is responsible for the loss of exon 7 inclusion in SMN2 mRNAs is a C to T transition at position 6 of exon 7. This C6T mutation disrupts an exonic splicing enhancer (ESE) that binds SRSF1 and generates an exonic splicing silencer (ESS) (Lorson *et al.*, 1999; Monani *et al.*, 1999; Cartegni & Krainer, 2002; Cartegni *et al.*, 2006). This silencer sequence has been shown to bind the repressor protein hnRNP A1 (Kashima & Manley, 2003; Kashima *et al.*, 2007) and has also been suggested to bind Sam68 (Pedrotti *et al.*, 2010), a regulatory protein that has been shown to interact with hnRNP A1 (Paronetto *et al.*, 2007). GGA anneals over this silencer site, but merely blocking this site with the annealing region alone provides only a small increase in SMN2 exon 7 inclusion compared to that seen with GGA (oligonucleotides 1 (GGA) and 8 (NT) in (Owen *et al.*, 2011)). It was therefore interesting to further investigate the bifunctional nature of GGA and how this affects its mechanism of action.

In order to investigate the potential role of the GGA annealing region (NT) to block Sam68 binding, single molecule methods were used to look at eGFP-Sam68 binding to SMN2 pre-mRNA in the presence and absence of the NT oligonucleotide (Figure 4.35). Although further controls to determine the ratio of eGFP labelled Sam68 to unlabeled endogenous Sam68 would be necessary to determine the exact number of Sam68

DISCUSSION

proteins bound per RNA, a significant reduction in the number of labelled Sam68 molecules bound per RNA was observed in the presence of the NT oligonucleotide. However, the overall level of SMN2 RNA molecules with any number of Sam68 proteins bound increased slightly from 26% to 34% when NT was used, potentially due to the disruption of larger aggregates of Sam68 by NT causing increased availability of Sam68 proteins. These results clearly indicate that the annealing region of GGA has the ability to alter Sam68 binding to an SMN2 pre-mRNA construct (Appendix 8.1.1). This finding is particularly interesting in light of other results obtained by colleague Dr L. Smith, which show that the NT oligonucleotide increases binding of U2AF65 to the SMN2 RNA (Smith *et al.*, manuscript submitted; Smith, 2012). This makes sense as the C6T mutation in SMN2 has been shown to reduce U2AF65 binding (De Araújo *et al.*, 2009). Sam68 has been found to antagonize U2AF65 binding in a different system, so it is not unreasonable to suggest that disruption of Sam68 binding would promote U2AF65 binding (Paronetto *et al.*, 2011).

The enhancer tail domain of GGA, which was shown to bind the activator protein SRSF1 (Section 7.3), adds significantly to the stimulatory effect on SMN2 exon 7 inclusion seen with the just the annealing region (Owen *et al.*, 2011). Release of the tail domain by photocleavage at different timepoints during an *in vitro* SMN2 splicing reaction resulted in progressively reduced functionality related to the time of cleavage (Figure 4.34). These results clearly indicate that the enhancer tail domain of GGA has a vital mechanistic role that works in addition to the silencer-blocking action of the annealing domain, as expected.

7.5 Overview of Proposed Mechanism of Action of GGA

The exon 7 C6T transition between SMN1 and SMN2, which destroys a binding site for SRSF1 and creates a binding site for hnRNP A1 and/or Sam68, has been shown to result in decreased U2AF65 binding as well as U2 snRNP recruitment (De Araújo *et al.*, 2009). Blocking of this mutation site with the annealing domain of GGA, results in recovery of U2AF65 binding. The enhancer tail domain, which binds SRSF1, replaces the disrupted SRSF1-binding ESE. This has been shown by colleague Dr L. Smith to result in stimulation/stabilization of U2 snRNP binding (Smith *et al.*, manuscript submitted; Smith, 2012). Therefore, the entire GGA oligonucleotide functions in two ways: (1) blocking the binding of repressor proteins hnRNP A1 and Sam68 through the annealing domain, allowing for U2AF65 binding, and (2) stimulating and/or stabilizing the binding of the U2 snRNP through contacts made by SRSF1, which binds to the enhancer tail domain. These findings can now be used to make more general assumptions about TOES. For example, they show the importance of removing U2AF65 repression, when applicable, by blocking the relevant silencer site. Also, they indicate that in cases where weak U2 snRNP binding is limiting, the enhancer tail domain of a targeted oligonucleotide enhancer of splicing would be sufficient to stimulate splicing, without the need to block a repressor site with the annealing region. Now that an outline mechanism of action of GGA has been elucidated, it provides a better guideline for the future design of therapeutic bifunctional oligonucleotides targeting splicing.

8 APPENDIX

8.1 Splicing Template Sequences

8.1.1 SMN2

β -GLOBIN Exon 2 (1-226 nt)

AGGGCTGCTGGTTGTCTACCCATGGACCCAGAGGTTCTTCGAGTCCTTTGGGGACCTGTC
CTCTGCAAATGCTGTTATGAACAATCCTAAGGTGAAGGCTCATGGCAAGAAGGTGCTGGC
TGCCTTCAGTGAGGGTCTGAGTCACCTGGACAACCTCAAAGGCACCTTTGCTAAGCTGAG
TGAAGTGCACCTGTGACAAGCTGCACGTGGATCCTGAGAAGCTTCAGG

END β -GLOBIN Exon 2/ BEGIN β -GLOBIN Intron 2 (226-293 nt)

GTGAGTTTGGGGACCCTTGATTGTTCTTTCTTTTTTCGCTATTGTAAAATTCATGTTATATGGTC
GAC

BEGIN SMN2 Intron 6 (293-450 nt)

AGACTATCAACTTAATTTCTGATCATATTTTGTGAATAAAAATAAGTAAAATGTCTT
GTGAAACAAAATGCTTTTTAACATCCATATAAAGCTATCTATATATAGCTATCTATATCT
ATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTACAG

BEGIN SMN2 Exon 7 (450-503 nt)

GGTTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATCAGGA

END SMN Exon 7/ BEGIN SMN2 Intron 7 (503-784 nt)

GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGT
AAAACCTTATGGTTTGTGGAAAACAAATGTTTTTGAACATTTAAAAAGTTCAGATGT
TAGAAAGTTGAAAGGTTAATGTAAAACAATCAATATTAAAGAATTTTGATGCCAAAACCTA
TTAGATAAAAAGGTTAATCTACATCCCTACTAGAATTCTCATACTTAACTGGTTGGTTGTG

APPENDIX

TGGAAGAAACATACTTTACAATAAAGAGCTTTAGGATATGATGCCATTTTATATCACGT
CGAC

END SMN Intron 7/BEGIN remainder of β -GLOBIN Intron 2 (784-822 nt)

TCTGCTAACCATGTCATGCCTTCTTCTTTTCCTACAG

BEGIN β -GLOBIN Exon 3 (822-871 nt)

CTCCTGGGCAACCGTGCTGGGTATTGGGCTGGCTCAACAATTTTGGCAG

END β -GLOBIN Exon 3 (871-878 nt)

GTAAGTT

8.1.2 SMN2 Exon 7 'Exon Defined' Construct

T7 promoter

AAATTAATACGACTCACTATAGGG

SMN2 Intron 6

ATATAGCTATTTTTTTAACTTCCTTTATTTTCCTTACAG

SMN2 Exon 7

GGTTTTAGACAAAATCAAAAAGAAGGAAGGTGCTCACATTCCTTAAATCAGGA

SMN2 Intron 7

GTAAGTCTGCCAGCATTATGAAAGTGAATCTTACTTTTGT

8.1.3 β -globin

β -GLOBIN Exon 2 (1-225 nt)

GGGCTGCTGGTTGTCTACCCATGGACCCAGAGGTTCTTCGAGTCCTTTGGGGACCTGTCCTCT
GCAAAATGCTGTTATGAACAATCCTAAGGTGAAGGCTCATGGCAAGAAGGTGCTGGCTGCCTC
AGTGAGGGTCTGAGTCACCTGGACAACCTCAAAGGCACCTTTGCTAAGCTGAGTGAAGTGCAC
TGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGG

APPENDIX

β-GLOBIN Intron 2 (225-331 nt)

GTGAGTTTGGGGACCCTTGATTGTTCTTTCTTTTTCGCTATTGTAAAATTCATGTTATATGGTCG
ACTCTGCTAACCATGTTTCATGCCTTCTTCTTTTTCTACAG

β-GLOBIN Exon 3 (331 -385 nt)

CTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGCAGGTAAGTT

8.1.4 β-globin + GGA anneal site

β-GLOBIN Exon 2 (1-225 nt)

GGGCTGCTGGTTGTCTACCCATGGACCCAGAGGTTCTTCGAGTCCTTTGGGGACCTGTCCTCT
GCAAATGCTGTTATGAACAATCCTAAGGTGAAGGCTCATGGCAAGAAGGTGCTGGCTGCCTTC
AGTGAGGGTCTGAGTCACCTGGACAACCTCAAAGGCACCTTTGCTAAGCTGAGTGAAGTGCAC
TGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGG

β-GLOBIN Intron 2 (225-331 nt)

GTGAGTTTGGGGACCCTTGATTGTTCTTTCTTTTTCGCTATTGTAAAATTCATGTTATATGGTCG
ACTCTGCTAACCATGTTTCATGCCTTCTTCTTTTTCTACAG

β-GLOBIN Exon 3 (331-376 nt)

CTCCTGGGCAACGTGCTGGTTATTGTGCTGTCTCATCATTTTGGC

GGA tail anneal site (376-396)

CCTCCGTCCTCCGTCCTCCT

8.1.5 Adenovirus WT (A2)

Note: Alternative 5' splice sites indicated in bold (GGG/GTGAGT). Intron sequence indicated in italics. (O'Mullane & Eperon, 1998; Lewis *et al.*, 2012)

T7 Promoter

AAATTAATACGACTCACTATAGGG

APPENDIX

Remainder of Construct

CATCGCTGTCTGCGAGGGCCAGCTGTT**GGGGTGAGT**ACTCCCTCTCAAAGCGGGCATGACT
TCTGCGCTAAGATTGTCAGTTTCAAAAACGAGGAGGATTTGATATTCACCAGCTGTT**GGGGT**
GAGTCCTTTGAGGGTGGCCGCGTCCATCTGGTCAGAAAAGACAATCTTTTTGTTGTCAAGCTT
GCTGCACGTCTAGGGCGCAGTAGTCCAGGGTTTCCTTGATGATGTCATACTTATCCTGTCCCTT
TTTTTCCACAGCTCGCGGTTGAGGACAACTCTTCGCGGTCTTCCAGTACTCTTGGATC

8.1.6 Adenovirus + ESE (A3)

Note: Alternative 5' splice sites indicated in bold (GGG/GTGAGT). Intron sequence indicated in italics. (O'Mullane & Eperon, 1998; Lewis *et al.*, 2012)

T7 Promoter

AAATTAATACGACTCACTATAGGG

ESE

AGGAGGACGGAGGACGGAGGA

Remainder of Construct

CATCGCTGTCTGCGAGGGCCAGCTGTT**GGGGTGAGT**ACTCCCTCTCAAAGCGGGCATGACT
TCTGCGCTAAGATTGTCAGTTTCAAAAACGAGGAGGATTTGATATTCACCAGCTGTT**GGGGT**
GAGTCCTTTGAGGGTGGCCGCGTCCATCTGGTCAGAAAAGACAATCTTTTTGTTGTCAAGCTT
GCTGCACGTCTAGGGCGCAGTAGTCCAGGGTTTCCTTGATGATGTCATACTTATCCTGTCCCTT
TTTTTCCACAGCTCGCGGTTGAGGACAACTCTTCGCGGTCTTCCAGTACTCTTGGATC

8.2 Oligonucleotide Sequences

Note: Phosphorothioate modified bases are indicated with a *. Bases with 2'OMethyl modification are indicated in bold. Locked nucleic acid (LNA) bases are indicated with a subscripted L.

8.2.1 SMN2 Splice-Altering RNA Oligonucleotides

GGA

A*G*G*A*G*GACGGAGGACGGAGGACAGAU*U*U*U*G*UCUAA*A*A*C

GGA-O

AGGAGGACGGAGGACGGAGGACAGAUUUUGUCUAAAAC

GGA-L/O

A_LG_LG_LA_LG_LGACGGAGGACGGAGGACAG_LAU_LUU_LUG_LUC_LUA_LAA_LAC

GGA-S

A*G*G*A*G*GACGGAGGACGGAGGACAGAU*U*U*U*G*UCUAA*A*A*C*

GGA-NS

AGGAGGACGGAGGACGGAGGACAGAUUUUGUCUAAAAC

GGA PC Bio (Synthesized by DNA Technology)

A*G*G*A*G*GACGGAGGACGGAGGACA – Photocleavable linker –
GAU*U*U*U*G*UCUAA*A*A*C – Biotin

TO

A*G*G*A*G*GACGGAGGACGGAGGAC*A*

O-TO

AGGAGGACGGAGGACGGAGGACA

NT

GAU*U*U*U*G*UCUAA*A*A*C

O-NT

GAUUUUGUCUAAAAC

APPENDIX

AAC

A*A*C*C*AGACGACAGACGAAAGAU*U*U*U*G*UCUAA*A*A*C

8.2.2 Oligonucleotides for SMN2 Fe-BABE Experiments

Note: All thiol modifying groups were added to the 5' end of the oligonucleotides. See Figure 6.1 for chemical structures of the protected/deprotected thiol modifier and Fe-BABE groups.

SH-GGA-O/Fe-BABE GGA-O

5' Mod. - **AGGAGGACGGAGGACGGAGGACAGAUUUUUGUCUAAAAC**

SH-NET-O/Fe-BABE NET-O

5' Mod. - **ACUUAUCGUUGAUUAUCAGUCAGGAUUUUGUCUAAAAC**

SH-GGA/Fe-BABE GGA

5' Mod. - **A*G*G*A*G*GACGGAGGACGGAGGACAGAU*U*U*U*G*UCUAA*A*A*C**

SH-NT/Fe-BABE NT

5' Mod. - **GAU*U*U*U*G*UCUAA*A*A*C**

8.2.3 Reverse Transcription Primers

GGA DNA Primer

GTTTTAGACAAAATC

SMN2 Primer 1

GTACAATATACCAGCTGTCTG

SMN2 Primer 2

TTTTCAAGTCTACAATCTTTCAACT

APPENDIX

SMN2 Primer 4

CGGTCGTAATACTTTCACTTAG

SMN2 Primer 5

ACCCGTTGGCACGACCCATAA

8.2.4 Other Oligonucleotides

β -globin 5' Cy5

Cy5-UA_LGA_LCA_LAC_LCA_LGC_LAG_LCC_L-Biotin

α -U1 bio

GCCAGGUAAGUAU-Biotin

α -U6 oligo

CUGUGUAUCGUUCCAAUUUU

BIBLIOGRAPHY

Akdim, F., Tribble, D.L., Flaim, J.D., Yu, R., Su, J., Geary, R.S., Baker, B.F., Fuhr, R., Wedel, M.K., Kastelein, J.J.P., 2011. Efficacy of apolipoprotein B synthesis inhibition in subjects with mild-to-moderate hyperlipidaemia. *European Heart Journal*. **32**, 2650-2659.

Alekseyenko, A.V., Kim, N., Lee, C.J., 2007. Global analysis of exon creation versus loss and the role of alternative splicing in 17 vertebrate genomes. *RNA*. **13**, 661-670.

Amariglio, N., Hirshberg, A., Scheithauer, B.W., Cohen, Y., Loewenthal, R., Trakhtenbrot, L., Paz, N., Koren-Michowitz, M., Waldman, D., Leider-Trejo, L., Toren, A., Constantini, S., Rechavi, G., 2009. Donor-derived brain tumor following neural stem cell transplantation in an ataxia telangiectasia patient. *PLoS Medicine*. **6**, 0221-0231.

Andersson, R., Enroth, S., Rada-Iglesias, A., Wadelius, C., Komorowski, J., 2009. Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Research*. **19**, 1732-1741.

Armas, P., Nasif, S., Calcaterra, N.B., 2008. Cellular nucleic acid binding protein binds G-rich single-stranded nucleic acids and may function as a nucleic acid chaperone. *Journal of Cellular Biochemistry*. **103**, 1013-1036.

Avila, A.M., Burnett, B.G., Taye, A.A., Gabanella, F., Knight, M.A., Hartenstein, P., Cizman, Z., Di Prospero, N.A., Pellizzoni, L., Fischbeck, K.H., Sumner, C.J., 2007. Trichostatin A increases SMN expression and survival in a mouse model of spinal muscular atrophy. *Journal of Clinical Investigation*. **117**, 659-671.

Azzouz, M., Le, T., Ralph, G.S., Walmsley, L., Monani, U.R., Lee, D.C.P., Wilkes, F., Mitrophanous, K.A., Kingsman, S.M., Burghes, A.H.M., Mazarakis, N.D., 2004. Lentivector-mediated SMN replacement in a mouse model of spinal muscular atrophy. *Journal of Clinical Investigation*. **114**, 1726-1731.

Baichoo, N. & Heyduk, T., 1997. Mapping conformational changes in a protein: Application of a protein footprinting technique to cAMP-induced conformational changes in cAMP receptor protein. *Biochemistry*. **36**, 10830-10836.

Banerjee, H., Rahn, A., Davis, W., Singh, R., 2003. Sex lethal and U2 small nuclear ribonucleoprotein auxiliary factor (U2AF65) recognize polypyrimidine tracts using multiple modes of binding. *RNA*. **9**, 88-99.

Banerjee, H., Rahn, A., Gawande, B., Guth, S., Valcárcel, J., Singh, R., 2004. The conserved RNA recognition motif 3 of U2 snRNA auxiliary factor (U2AF65) is essential in vivo but dispensable for activity in vitro. *RNA*. **10**, 240-253.

BIBLIOGRAPHY

- Bashkirov, V.I., Scherthan, H., Solinger, J.A., Buerstedde, J.-., Heyer, W.-., 1997. A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. *Journal of Cell Biology*. **136**, 761-773.
- Bates, P.J., Kahlon, J.B., Thomas, S.D., Trent, J.O., Miller, D.M., 1999. Antiproliferative activity of G-rich oligonucleotides correlates with protein binding. *Journal of Biological Chemistry*. **274**, 26369-26377.
- Baughan, T., Shababi, M., Coady, T.H., Dickson, A.M., Tullis, G.E., Lorson, C.L., 2006. Stimulating Full-Length SMN2 Expression by Delivering Bifunctional RNAs via a Viral Vector. *Molecular Therapy*. **14**, 54-62.
- Baughan, T.D., Dickson, A., Osman, E.Y., Lorson, C.L., 2009. Delivery of bifunctional RNAs that target an intronic repressor and increase SMN levels in an animal model of spinal muscular atrophy. *Human Molecular Genetics*. **18**, 1600-1611.
- Beaudoin, J.-. & Perreault, J.-., 2010. 5'-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Research*. **38**, 7022-7036.
- Beaudoing, E., Freier, S., Wyatt, J.R., Claverie, J.-., Gautheret, D., 2000. Patterns of variant polyadenylation signal usage in human genes. *Genome Research*. **10**, 1001-1010.
- Bebbee, T.W., Dominguez, C.E., Chandler, D.S., 2012. Mouse models of SMA: Tools for disease characterization and therapeutic development. *Human Genetics*. **131**, 1277-1293.
- Berget, S.M., 1995. Exon recognition in vertebrate splicing. *Journal of Biological Chemistry*. **270**, 2411-2414.
- Berget, S.M., Moore, C., Sharp, P.A., 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences of the United States of America*. **74**, 3171-3175.
- Berglund, J.A., Abovich, N., Rosbash, M., 1998. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes and Development*. **12**, 858-867.
- Bochman, M.L., Paeschke, K., Zakian, V.A., 2012. DNA secondary structures: Stability and function of G-quadruplex structures. *Nature Reviews Genetics*. **13**, 770-780.
- Borgognone, M., Armas, P., Calcaterra, N.B., 2010. Cellular nucleic-acid-binding protein, a transcriptional enhancer of c-Myc, promotes the formation of parallel G-quadruplexes. *Biochemical Journal*. **428**, 491-498.

BIBLIOGRAPHY

Bourgeois, C.F., Popielarz, M., Hildwein, G., Stevenin, J., 1999. Identification of a bidirectional splicing enhancer: Differential involvement of SR proteins in 5' or 3' splice site activation. *Molecular and Cellular Biology*. **19**, 7347-7356.

Boutz, P.L., Stoilov, P., Li, Q., Lin, C.-., Chawla, G., Ostrow, K., Shiue, L., Ares Jr., M., Black, D.L., 2007. A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes and Development*. **21**, 1636-1652.

Bringmann, P., Appel, B., Rinke, J., Reuter, R., Theissen, H., Lührmann, R., 1984. Evidence for the existence of snRNAs U4 and U6 in a single ribonucleoprotein complex and for their association by intermolecular base pairing. *EMBO Journal*. **3**, 1357-1363.

Burghes, A.H.M. & Beattie, C.E., 2009. Spinal muscular atrophy: Why do low levels of survival motor neuron protein make motor neurons sick? *Nature Reviews Neuroscience*. **10**, 597-609.

Caputi, M. & Zahler, A.M., 2001. Determination of the RNA Binding Specificity of the Heterogeneous Nuclear Ribonucleoprotein (hnRNP) H/H'/F/2H9 Family. *Journal of Biological Chemistry*. **276**, 43850-43859.

Carrel, T.L., McWhorter, M.L., Workman, E., Zhang, H., Wolstencroft, E.C., Lorson, C., Bassell, G.J., Burghes, A.H.M., Beattie, C.E., 2006. Survival motor neuron function in motor axons is independent of functions required for small nuclear ribonucleoprotein biogenesis. *Journal of Neuroscience*. **26**, 11014-11022.

Cartegni, L., Hastings, M.L., Calarco, J.A., De Stanchina, E., Krainer, A.R., 2006. Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. *American Journal of Human Genetics*. **78**, 63-77.

Cartegni, L. & Krainer, A.R., 2003. Correction of disease-associated exon skipping by synthetic exon-specific activators. *Nature Structural Biology*. **10**, 120-125.

Cartegni, L. & Krainer, A.R., 2002. Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN. *Nature Genetics*. **30**, 377-384.

Cartegni, L., Wang, J., Zhu, Z., Zhang, M.Q., Krainer, A.R., 2003. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Research*. **31**, 3568-3571.

Chen, M. & Manley, J.L., 2009. Mechanisms of alternative splicing regulation: Insights from molecular and genomics approaches. *Nature Reviews Molecular Cell Biology*. **10**, 741-754.

Cho, S., Hoang, A., Sinha, R., Zhong, X.-., Fu, X.-., Krainer, A.R., Ghosh, G., 2011. Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K

BIBLIOGRAPHY

snRNP protein determines early spliceosome assembly. *Proceedings of the National Academy of Sciences of the United States of America*. **108**, 8233-8238.

Chou, M.-., Rooke, N., Turck, C.W., Black, D.L., 1999. hnRNP H is a component of a splicing enhancer complex that activates a c-src alternative exon in neuronal cells. *Molecular and Cellular Biology*. **19**, 69-77.

Chow, L.T., Gelinas, R.E., Broker, T.R., Roberts, R.J., 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*. **12**, 1-8.

Cirak, S., Arechavala-Gomez, V., Guglieri, M., Feng, L., Torelli, S., Anthony, K., Abbs, S., Garralda, M.E., Bourke, J., Wells, D.J., Dickson, G., Wood, M.J., Wilton, S.D., Straub, V., Kole, R., Shrewsbury, S.B., Sewry, C., Morgan, J.E., Bushby, K., Muntoni, F., 2011. Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: An open-label, phase 2, dose-escalation study. *The Lancet*. **378**, 595-605.

Coady, T.H. & Lorson, C.L., 2011. SMN in spinal muscular atrophy and snRNP biogenesis. *Wiley Interdisciplinary Reviews: RNA*. **2**, 546-564.

Colbourne, J.K., Pfrender, M.E., Gilbert, D., Thomas, W.K., Tucker, A., Oakley, T.H., Tokishita, S., Aerts, A., Arnold, G.J., Basu, M.K., Bauer, D.J., Cáceres, C.E., Carmel, L., Casola, C., Choi, J.-., Detter, J.C., Dong, Q., Dusheyko, S., Eads, B.D., Fröhlich, T., Geiler-Samerotte, K.A., Gerlach, D., Hatcher, P., Jogdeo, S., Krijgsveld, J., Kriventseva, E.V., Kültz, D., Laforsch, C., Lindquist, E., Lopez, J., Manak, J.R., Muller, J., Pangilinan, J., Patwardhan, R.P., Pitluck, S., Pritham, E.J., Rechtsteiner, A., Rho, M., Rogozin, I.B., Sakarya, O., Salamov, A., Schaack, S., Shapiro, H., Shiga, Y., Skalitzky, C., Smith, Z., Suvorov, A., Sung, W., Tang, Z., Tsuchiya, D., Tu, H., Vos, H., Wang, M., Wolf, Y.I., Yamagata, H., Yamada, T., Ye, Y., Shaw, J.R., Andrews, J., Crease, T.J., Tang, H., Lucas, S.M., Robertson, H.M., Bork, P., Koonin, E.V., Zdobnov, E.M., Grigoriev, I.V., Lynch, M., Boore, J.L., 2011. The ecoresponsive genome of *Daphnia pulex*. *Science*. **331**, 555-561.

Coolidge, C.J., Seely, R.J., Patton, J.G., 1997. Functional analysis of the polypyrimidine tract in pre-mRNA splicing. *Nucleic Acids Research*. **25**, 888-895.

Cooper, T.A., Wan, L., Dreyfuss, G., 2009. RNA and Disease. *Cell*. **136**, 777-793.

Covert, D.D., Le, T.T., McAndrew, P.E., Strasswimmer, J., Crawford, T.O., Mendell, J.R., Coulson, S.E., Androphy, E.J., Prior, T.W., Burghes, A.H.M., 1997. The survival motor neuron protein in spinal muscular atrophy. *Human Molecular Genetics*. **6**, 1205-1214.

Corti, S., Nizzardo, M., Nardini, M., Donadoni, C., Salani, S., Ronchi, D., Simone, C., Falcone, M., Papadimitriou, D., Locatelli, F., Mezzina, N., Gianni, F., Bresolin, N., Comi, G.P., 2010. Embryonic stem cell-derived neural stem cells improve spinal muscular atrophy phenotype in mice. *Brain*. **133**, 465-481.

BIBLIOGRAPHY

- Creacy, S.D., Routh, E.D., Iwamoto, F., Nagamine, Y., Akman, S.A., Vaughn, J.P., 2008. G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. *Journal of Biological Chemistry*. **283**, 34626-34634.
- Crispino, J.D., Blencowe, B.J., Sharp, P.A., 1994. Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science*. **265**, 1866-1869.
- Crispino, J.D. & Sharp, P.A., 1995. A U6 snRNA:pre-mRNA interaction can be rate-limiting for U1-independent splicing. *Genes and Development*. **9**, 2314-2323.
- Damgaard, C.K., Tange, T.Ø, Kjems, J., 2002. HnRNP A1 controls HIV-1 mRNA splicing through cooperative binding to intron and exon splicing silencers in the context of a conserved secondary structure. *RNA*. **8**, 1401-1415.
- Das, R. & Reed, R., 1999. Resolution of the mammalian E complex and the ATP-dependent spliceosomal complexes on native agarose mini-gels. *RNA*. **5**, 1504-1508.
- De Araújo, M.M., Bonnal, S., Hastings, M.L., Krainer, A.R., Valcárcel, J., 2009. Differential 3' splice site recognition of SMN1 and SMN2 transcripts by U2AF and U2 snRNP. *RNA*. **15**, 515-523.
- De Conti, L., Baralle, M., Buratti, E., 2013. Exon and intron definition in pre-mRNA splicing. *Wiley Interdisciplinary Reviews: RNA*. **4**, 49-60.
- Deleavey, G. & Damha, M., 2012. Designing Chemically Modified Oligonucleotides for Targeted Gene Silencing. *Chemistry & Biology*. **19**, 937-954.
- Didiot, M.-., Tian, Z., Schaeffer, C., Subramanian, M., Mandel, J.-., Moine, H., 2008. The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer. *Nucleic Acids Research*. **36**, 4902-4912.
- Dominguez, C. & Allain, F.H.-., 2006. NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: A novel mode of RNA recognition. *Nucleic Acids Research*. **34**, 3634-3645.
- Dominguez, C., Fiset, J.-., Chabot, B., Allain, F.H.-., 2010. Structural basis of G-tract recognition and encaging by hnRNP F quasi-RRMs. *Nature Structural and Molecular Biology*. **17**, 853-861.
- Dominguez, E., Marais, T., Chatauret, N., Benkhelifa-Ziyyat, S., Duque, S., Ravassard, P., Carcenac, R., Astord, S., de Moura, A.P., Voit, T., Barkats, M., 2011. Intravenous scAAV9 delivery of a codon-optimized SMN1 sequence rescues SMA mice. *Human Molecular Genetics*. **20**, 681-693.

BIBLIOGRAPHY

Dönmez, G., Hartmuth, K., Kastner, B., Will, C.L., Lührmann, R., 2007. The 5' End of U2 snRNA Is in Close Proximity to U1 and Functional Sites of the Pre-mRNA in Early Spliceosomal Complexes. *Molecular Cell*. **25**, 399-411.

Donnelly, E.M. & Boulis, N.M., 2012. Update on gene and stem cell therapy approaches for spinal muscular atrophy. *Expert Opinion on Biological Therapy*. **12**, 1463-1471.

Dreyer, G.B. & Dervan, P.B., 1985. Sequence-specific cleavage of single-stranded DNA: Oligodeoxynucleotide-EDTA·Fe(II). *Proceedings of the National Academy of Sciences of the United States of America*. **82**, 968-972.

Du, H. & Rosbash, M., 2002. The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing. *Nature*. **419**, 86-90.

Düchler, M., 2012. G-quadruplexes: Targets and tools in anticancer drug design. *Journal of Drug Targeting*. **20**, 389-400.

Dujardin, G., Lafaille, C., Petrillo, E., Buggiano, V., Gómez Acuña, L.I., Fiszbein, A., Godoy Herz, M.A., Nieto Moreno, N., Muñoz, M.J., Alló, M., Schor, I.E., Kornblihtt, A.R., 2013. Transcriptional elongation and alternative splicing. *Biochimica Et Biophysica Acta - Gene Regulatory Mechanisms*. **1829**, 134-140.

Eddy, J. & Maizels, N., 2006. Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Research*. **34**, 3887-3896.

Eperon, I., 2012. Antisense therapeutics: New ways to nudge splicing. *Nature Chemical Biology*. **8**, 507-508.

Eperon, I.C., Ireland, D.C., Smith, R.A., Mayeda, A., Krainer, A.R., 1993. Pathways for selection of 5' splice sites by U1 snRNPs and SF2/ASF. *EMBO Journal*. **12**, 3607-3617.

Eperon, I.C., Makarova, O.V., Mayeda, A., Munroe, S.H., Caceres, J.F., Hayward, D.G., 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**, 8303-8318.

Eperon, L.P., Estibeiro, J.P., Eperon, I.C., 1986. The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA. *Nature*. **324**, 280-282.

Eperon, L.P., Graham, I.R., Griffiths, A.D., Eperon, I.C., 1988. Effects of RNA secondary structure on alternative splicing of Pre-mRNA: Is folding limited to a region behind the transcribing RNA polymerase? *Cell*. **54**, 393-401.

Erkelenz, S., Mueller, W.F., Evans, M.S., Busch, A., Schöneweis, K., Hertel, K.J., Schaal, H., 2013. Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms. *RNA*. **19**, 96-102.

BIBLIOGRAPHY

- Expert-Bezançon, A., Sureau, A., Durosay, P., Salesse, R., Groeneveld, H., Lecaer, J.P., Marie, J., 2004. hnRNP A1 and the SR proteins ASF/SF2 and SC35 have antagonistic functions in splicing of β -tropomyosin exon 6B. *Journal of Biological Chemistry*. **279**, 38249-38259.
- Fairbrother, W.G., Yeh, R.-., Sharp, P.A., Burge, C.B., 2002. Predictive identification of exonic splicing enhancers in human genes. *Science*. **297**, 1007-1013.
- Fan, L., Lagiseti, C., Edwards, C.C., Webb, T.R., Potter, P.M., 2011. Sudemycins, novel small molecule analogues of FR901464, induce alternative gene splicing. *ACS Chemical Biology*. **6**, 582-589.
- Feldkötter, M., Schwarzer, V., Wirth, R., Wienker, T.F., Wirth, B., 2002. Quantitative analyses of SMN1 and SMN2 based on real-time lightcycler PCR: Fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *American Journal of Human Genetics*. **70**, 358-368.
- Fletcher, T.M., Sun, D., Salazar, M., Hurley, L.H., 1998. Effect of DNA secondary structure on human telomerase activity. *Biochemistry*. **37**, 5536-5541.
- Flouriot, G., Brand, H., Seraphin, B., Gannon, F., 2002. Natural trans-spliced mRNAs are generated from the human estrogen receptor- α (hER α) gene. *Journal of Biological Chemistry*. **277**, 26244-26251.
- Fortner, D.M., Troy, R.G., Brow, D.A., 1994. A stem/loop in U6 RNA defines a conformational switch required for pre- mRNA splicing. *Genes and Development*. **8**, 221-233.
- Foust, K.D., Wang, X., McGovern, V.L., Braun, L., Bevan, A.K., Haidet, A.M., Le, T.T., Morales, P.R., Rich, M.M., Burghes, A.H.M., Kaspar, B.K., 2010. Rescue of the spinal muscular atrophy phenotype in a mouse model by early postnatal delivery of SMN. *Nature Biotechnology*. **28**, 271-274.
- Fox-Walsh, K.L., Dou, Y., Lam, B.J., Hung, S.-., Baldi, P.F., Hertel, K.J., 2005. The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *Proceedings of the National Academy of Sciences of the United States of America*. **102**, 16176-16181.
- Fry, M., 2007. Tetraplex DNA and its interacting proteins. *Frontiers in Bioscience*. **12**, 4336-4351.
- Fukuhara, T., Hosoya, T., Shimizu, S., Sumi, K., Oshiro, T., Yoshinaka, Y., Suzuki, M., Yamamoto, N., Herzenberg, L.A., Herzenberg, L.A., Hagiwara, M., 2006. Utilization of host SR protein kinases and RNA-splicing machinery during viral replication. *Proceedings of the National Academy of Sciences of the United States of America*. **103**, 11329-11333.

BIBLIOGRAPHY

- Fukumura, K., Taniguchi, I., Sakamoto, H., Ohno, M., Inoue, K., 2009. U1-independent pre-mRNA splicing contributes to the regulation of alternative splicing. *Nucleic Acids Research*. **37**, 1907-1914.
- Gabut, M., Miné, M., Marsac, C., Brivet, M., Tazi, J., Soret, J., 2005. The SR protein SC35 is responsible for aberrant splicing of the E1 α pyruvate dehydrogenase mRNA in a case of mental retardation with lactic acidosis. *Molecular and Cellular Biology*. **25**, 3286-3294.
- Garbes, L., Riessland, M., Hölker, I., Heller, R., Hauke, J., Tränkle, C., Coras, R., Blümcke, I., Hahnen, E., Wirth, B., 2009. LBH589 induces up to 10-fold SMN protein levels by several independent mechanisms and is effective even in cells from SMA patients non-responsive to valproate. *Human Molecular Genetics*. **18**, 3645-3658.
- Gaur, R.K., Valcárcel, J., Green, M.R., 1995. Sequential recognition of the pre-mRNA branch point by U2AF65 and a novel spliceosome-associated 28-kDa protein. *RNA*. **1**, 407-417.
- Geib, T. & Hertel, K.J., 2009. Restoration of full-length SMN promoted by adenoviral vectors expressing RNA antisense oligonucleotides embedded in U7 snRNAs. *PLoS One*. **4**, .
- GELLERT, M., LIPSETT, M.N., DAVIES, D.R., 1962. Helix formation by guanylic acid. *Proceedings of the National Academy of Sciences of the United States of America*. **48**, 2013-2018.
- Ghigna, C., De Toledo, M., Bonomi, S., Valacca, C., Gallo, S., Apicella, M., Eperon, I., Tazi, J., Biamonti, G., 2010. Pro-metastatic splicing of Ron proto-oncogene mRNA can be reversed: Therapeutic potential of bifunctional oligonucleotides and indole derivatives. *RNA Biology*. **7**, 495-503.
- Ghigna, C., Giordano, S., Shen, H., Benvenuto, F., Castiglioni, F., Comoglio, P., Green, M., Riva, S., Biamonti, G., 2005. Cell motility is controlled by SF2/ASF through alternative splicing of the Ron protooncogene. *Molecular Cell*. **20**, 881-890.
- Ghigna, C., Valacca, C., Biamonti, G., 2008. Alternative Splicing and Tumor Progression. *Current Genomics*. **9**, 556-570.
- Giri, B., Smaldino, P.J., Thys, R.G., Creacy, S.D., Routh, E.D., Hantgan, R.R., Lattmann, S., Nagamine, Y., Akman, S.A., Vaughn, J.P., 2011. G4 Resolvase 1 tightly binds and unwinds unimolecular G4-DNA. *Nucleic Acids Research*. **39**, 7161-7178.
- Goemans, N.M., Tulinius, M., Van Den Akker, J.T., Burm, B.E., Ekhardt, P.F., Heuvelmans, N., Holling, T., Janson, A.A., Platenburg, G.J., Sipkens, J.A., Sitsen, J.M.A., Aartsma-Rus, A., Van Ommen, G.-B., Buyse, G., Darin, N., Verschuuren, J.J., Campion, G.V., Kimpe, S.J.D., Van Deutekom, J.C., 2011. Systemic administration of PRO051 in Duchenne's muscular dystrophy. *New England Journal of Medicine*. **364**, 1513-1522.

BIBLIOGRAPHY

- Gomez, D., Lamarteleur, T., Lacroix, L., Mailliet, P., Mergny, J.-., Riou, J.-., 2004. Telomerase downregulation induced by the G-quadruplex ligand 12459 in A549 cells is mediated by hTERT RNA alternative splicing. *Nucleic Acids Research*. **32**, 371-379.
- González, V., Guo, K., Hurley, L., Sun, D., 2009. Identification and characterization of nucleolin as a c-myc G-quadruplex-binding protein. *Journal of Biological Chemistry*. **284**, 23622-23635.
- González, V. & Hurley, L.H., 2010. The C-terminus of nucleolin promotes the formation of the c-MYC G-quadruplex and inhibits c-MYC promoter activity. *Biochemistry*. **49**, 9706-9714.
- Görnemann, J., Kotovic, K.M., Hujer, K., Neugebauer, K.M., 2005. Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Molecular Cell*. **19**, 53-63.
- Grabowski, P.J., Padgett, R.A., Sharp, P.A., 1984. Messenger RNA splicing in vitro: An excised intervening sequence and a potential intermediate. *Cell*. **37**, 415-427.
- Graveley, B.R., 2005. Small molecule control of pre-mRNA splicing. *RNA*. **11**, 355-358.
- Graveley, B.R., Hertel, K.J., Maniatis, T.O.M., 2001. The role of U2AF³⁵ and U2AF⁶⁵ in enhancer-dependent splicing. *RNA*. **7**, 806-818.
- Graveley, B.R. & Maniatis, T., 1998. Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Molecular Cell*. **1**, 765-771.
- Greiner, D.P., Miyake, R., Moran, J.K., Jones, A.D., Negishi, T., Ishihama, A., Meares, C.F., 1997. Synthesis of the protein cutting reagent iron (S)-1-(p-bromoacetamidobenzyl)ethylenediaminetetraacetate and conjugation to cysteine side chains. *Bioconjugate Chemistry*. **8**, 44-48.
- Gros, J., Guédin, A., Mergny, J.-., Lacroix, L., 2008. G-quadruplex formation interferes with P1 helix formation in the RNA component of telomerase hTERC. *ChemBioChem*. **9**, 2075-2079.
- Hamilton, G. & Gillingwater, T.H., 2013. Spinal muscular atrophy: Going beyond the motor neuron. *Trends in Molecular Medicine*. **19**, 40-50.
- Han, K., Yeo, G., An, P., Burge, C.B., Grabowski, P.J., 2005. A combinatorial code for splicing silencing: UAGG and GGGG motifs. *PLoS Biology*. **3**, 0843-0860.
- Han, S.P., Tang, Y.H., Smith, R., 2010. Functional diversity of the hnRNPs: Past, present and perspectives. *Biochemical Journal*. **430**, 379-392.

BIBLIOGRAPHY

- Harrington, C., Lan, Y., Akman, S.A., 1997. The identification and characterization of a G4-DNA resolvase activity. *Journal of Biological Chemistry*. **272**, 24631-24636.
- Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., Barnes, I., Bignell, A., Boychenko, V., Hunt, T., Kay, M., Mukherjee, G., Rajan, J., Despacio-Reyes, G., Saunders, G., Steward, C., Harte, R., Lin, M., Howald, C., Tanzer, A., Derrien, T., Chrast, J., Walters, N., Balasubramanian, S., Pei, B., Tress, M., Rodriguez, J.M., Ezkurdia, I., Van Baren, J., Brent, M., Haussler, D., Kellis, M., Valencia, A., Reymond, A., Gerstein, M., Guigó, R., Hubbard, T.J., 2012. GENCODE: The reference human genome annotation for the ENCODE project. *Genome Research*. **22**, 1760-1774.
- Hashimoto, C. & Steitz, J.A., 1984. U4 and U6 RNAs coexist in a single small nuclear ribonucleoprotein particle. *Nucleic Acids Research*. **12**, 3283-3293.
- Hastings, M.L. & Krainer, A.R., 2001. Pre-mRNA splicing in the new millennium. *Current Opinion in Cell Biology*. **13**, 302-309.
- Hauke, J., Riessland, M., Lunke, S., Eyüpoglu, I.Y., Blümcke, I., El-osta, A., Wirth, B., Hahnen, E., 2009. Survival motor neuron gene 2 silencing by DNA methylation correlates with spinal muscular atrophy disease severity and can be bypassed by histone deacetylase inhibition. *Human Molecular Genetics*. **18**, 304-317.
- Hawkins, J.D., 1988. A survey on intron and exon lengths. *Nucleic Acids Research*. **16**, 9893-9908.
- Henscheid, K.L., Voelker, R.B., Berglund, J.A., 2008. Alternative modes of binding by U2AF65 at the polypyrimidine tract. *Biochemistry*. **47**, 449-459.
- Hermann, H., Fabrizio, P., Raker, V.A., Foulaki, K., Hornig, H., Brahm, H., Luhrmann, R., 1995. snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO Journal*. **14**, 2076-2088.
- Hodson, M.J., Hudson, A.J., Cherny, D., 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*. **40**, 6850-6862.
- Hofmann, Y., Lorson, C.L., Stamm, S., Androphy, E.J., Wirth, B., 2000. Htra2- β 1 stimulates an exonic splicing enhancer and can restore full-length SMN expression to survival motor neuron 2 (SMN2). *Proceedings of the National Academy of Sciences of the United States of America*. **97**, 9618-9623.
- Hong, W., Bennett, M., Xiao, Y., Kramer, R.F., Wang, C., Reed, R., 1997. Association of U2 snRNP with the spliceosomal complex E. *Nucleic Acids Research*. **25**, 354-361.
- Hong, X., Scofield, D.G., Lynch, M., 2006. Intron size, abundance, and distribution within untranslated regions of genes. *Molecular Biology and Evolution*. **23**, 2392-2404.

BIBLIOGRAPHY

- Hua, Y., Sahashi, K., Hung, G., Rigo, F., Passini, M.A., Bennett, C.F., Krainer, A.R., 2010. Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes and Development*. **24**, 1634-1644.
- Hua, Y., Sahashi, K., Rigo, F., Hung, G., Horev, G., Bennett, C.F., Krainer, A.R., 2011. Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model. *Nature*. **478**, 123-126.
- Hua, Y., Vickers, T.A., Baker, B.F., Bennett, C.F., Krainer, A.R., 2007. Enhancement of SMN2 exon 7 inclusion by antisense oligonucleotides targeting the exon. *PLoS Biology*. **5**, 729-744.
- Hua, Y., Vickers, T.A., Okunola, H.L., Bennett, C.F., Krainer, A.R., 2008. Antisense Masking of an hnRNP A1/A2 Intronic Splicing Silencer Corrects SMN2 Splicing in Transgenic Mice. *American Journal of Human Genetics*. **82**, 834-848.
- Huppert, J.L., 2010. Structure, location and interactions of G-quadruplexes. *FEBS Journal*. **277**, 3452-3458.
- Huppert, J.L. & Balasubramanian, S., 2007. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Research*. **35**, 406-413.
- Huppert, J.L. & Balasubramanian, S., 2005. Prevalence of quadruplexes in the human genome. *Nucleic Acids Research*. **33**, 2908-2916.
- Huppert, J.L., Bugaut, A., Kumari, S., Balasubramanian, S., 2008. G-quadruplexes: The beginning and end of UTRs. *Nucleic Acids Research*. **36**, 6260-6268.
- Huranová, M., Ivani, I., Benda, A., Poser, I., Brody, Y., Hof, M., Shav-Tal, Y., Neugebauer, K.M., Staněk, D., 2010. The differential interaction of snRNPs with pre-mRNA reveals splicing kinetics in living cells. *Journal of Cell Biology*. **191**, 75-86.
- Indig, F.E., Rybanska, I., Karmakar, P., Devulapalli, C., Fu, H., Carrier, F., Bohr, V.A., 2012. Nucleolin inhibits G4 oligonucleotide unwinding by werner helicase. *PLoS ONE*. **7**, .
- Jim, R., Gaffney, B.L., Wang, C., Jones, R.A., Breslauer, K.J., 1992. Thermodynamics and structure of a DNA tetraplex: A spectroscopic and calorimetric study of the tetramolecular complexes of d(TG3T) and d(TG3T2G3T). *Proceedings of the National Academy of Sciences of the United States of America*. **89**, 8832-8836.
- Jin, Y., Yang, Y., Zhang, P., 2011. New insights into RNA secondary structure in the alternative splicing of pre-mRNAs. *RNA Biology*. **8**, 450-457.
- Joseph, S. and Noller, H.F., 2000. [13] Directed hydroxyl radical probing using iron(II) tethered to RNA.

BIBLIOGRAPHY

Kaida, D., Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K., Watanabe, H., Kitahara, T., Yoshida, T., Nakajima, H., Tani, T., Horinouchi, S., Yoshida, M., 2007. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nature Chemical Biology*. **3**, 576-583.

Kaida, D., Schneider-Poetsch, T., Yoshida, M., 2012. Splicing in oncogenesis and tumor suppression. *Cancer Science*. **103**, 1611-1616.

Kan, J.L.C. & Green, M.R., 1999. Pre-mRNA splicing of IgM exons M1 and M2 is directed by a juxtaposed splicing enhancer and inhibitor. *Genes and Development*. **13**, 462-471.

Karni, R., De Stanchina, E., Lowe, S.W., Sinha, R., Mu, D., Krainer, A.R., 2007. The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nature Structural and Molecular Biology*. **14**, 185-193.

Kashima, T. & Manley, J.L., 2003. A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. *Nature Genetics*. **34**, 460-463.

Kashima, T., Rao, N., David, C.J., Manley, J.L., 2007. hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. *Human Molecular Genetics*. **16**, 3149-3159.

Kashima, T., Rao, N., Manley, J.L., 2007. An intronic element contributes to splicing repression in spinal muscular atrophy. *Proceedings of the National Academy of Sciences of the United States of America*. **104**, 3426-3431.

Kent, O.A. & MacMillan, A.M., 2002. Early organization of pre-mRNA during spliceosome assembly. *Nature Structural Biology*. **9**, 576-581.

Keren, H., Lev-Maor, G., Ast, G., 2010. Alternative splicing and evolution: Diversification, exon definition and function. *Nature Reviews Genetics*. **11**, 345-355.

Kernochan, L.E., Russo, M.L., Woodling, N.S., Huynh, T.N., Avila, A.M., Fischbeck, K.H., Sumner, C.J., 2005. The role of histone acetylation in SMN gene expression. *Human Molecular Genetics*. **14**, 1171-1182.

Kielkopf, C.L., Rodionova, N.A., Green, M.R., Burley, S.K., 2001. A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. *Cell*. **106**, 595-605.

Kikin, O., Zappala, Z., D'Antonio, L., Bagga, P.S., 2008. GRSD2 and GRS_UTRdb: Databases of quadruplex forming G-rich sequences in pre-mRNAs and mRNAs. *Nucleic Acids Research*. **36**, .

Kim, D.-., Kim, J.-., Park, M., Yeom, J.-., Go, H., Kim, S., Han, M.S., Lee, K., Bae, J., 2011. Modulation of biological processes in the nucleus by delivery of DNA oligonucleotides conjugated with gold nanoparticles. *Biomaterials*. **32**, 2593-2604.

BIBLIOGRAPHY

- Kim, E., Goren, A., Ast, G., 2008. Alternative splicing: Current perspectives. *BioEssays*. **30**, 38-47.
- Kim, E., Magen, A., Ast, G., 2007. Different levels of alternative splicing among eukaryotes. *Nucleic Acids Research*. **35**, 125-131.
- Kinali, M., Arechavala-Gomez, V., Feng, L., Cirak, S., Hunt, D., Adkin, C., Guglieri, M., Ashton, E., Abbs, S., Nihoyannopoulos, P., Garralda, M.E., Rutherford, M., Mcculley, C., Popplewell, L., Graham, I.R., Dickson, G., Wood, M.J., Wells, D.J., Wilton, S.D., Kole, R., Straub, V., Bushby, K., Sewry, C., Morgan, J.E., Muntoni, F., 2009. Local restoration of dystrophin expression with the morpholino oligomer AVI-4658 in Duchenne muscular dystrophy: a single-blind, placebo-controlled, dose-escalation, proof-of-concept study. *The Lancet Neurology*. **8**, 918-928.
- Kiss, T., 2004. Biogenesis of small nuclear RNPs. *Journal of Cell Science*. **117**, 5949-5951.
- Kling, J., 2010. Safety signal dampens reception for mipomersen antisense. *Nature Biotechnology*. **28**, 295-297.
- Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S., Ahringer, J., 2009. Differential chromatin marking of introns and expressed exons by H3K36me3. *Nature Genetics*. **41**, 376-381.
- Konarska, M.M., Grabowski, P.J., Padgett, R.A., Sharp, P.A., 1985. Characterization of the branch site in lariat RNAs produced by splicing of mRNA precursors. *Nature*. **313**, 552-557.
- Kong, D. & Yamori, T., 2012. JFCR39, a panel of 39 human cancer cell lines, and its application in the discovery and development of anticancer drugs. *Bioorganic and Medicinal Chemistry*. **20**, 1947-1951.
- Kornblihtt, A.R., De La Mata, M., Fededa, J.P., Muñoz, M.J., Nogués, G., 2004. Multiple links between transcription and splicing. *RNA*. **10**, 1489-1498.
- Kostadinov, R., Malhotra, N., Viotti, M., Shine, R., D'Antonio, L., Bagga, P., 2006. GRSDDB: a database of quadruplex forming G-rich sequences in alternatively processed mammalian pre-mRNA sequences. *Nucleic Acids Research*. **34**, .
- Kotake, Y., Sagane, K., Owa, T., Mimori-Kiyosue, Y., Shimizu, H., Uesugi, M., Ishihama, Y., Iwata, M., Mizui, Y., 2007. Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nature Chemical Biology*. **3**, 570-575.
- Kramer, A. & Utans, U., 1991. Three protein factors (SF1, SF3 and U2AF) function in pre-splicing complex formation in addition to snRNPs. *EMBO Journal*. **10**, 1503-1509.

BIBLIOGRAPHY

- Krawczak, M., Thomas, N.S.T., Hundrieser, B., Mort, M., Wittig, M., Hampe, J., Cooper, D.N., 2007. Single base-pair substitutions in exon-intron junctions of human genes: Nature, distribution, and consequences for mRNA splicing. *Human Mutation*. **28**, 150-158.
- Lane, A.N., Chaires, J.B., Gray, R.D., Trent, J.O., 2008. Stability and kinetics of G-quadruplex structures. *Nucleic Acids Research*. **36**, 5482-5515.
- Lang, K.M. & Spritz, R.A., 1983. RNA splice site selection: Evidence for a 5'→3' scanning model. *Science*. **220**, 1351-1355.
- Lavigueur, A., La Branche, H., Kornblihtt, A.R., Chabot, B., 1993. A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes and Development*. **7**, 2405-2417.
- Le, T.T., Coovert, D.D., Monani, U.R., Morris, G.E., Burghes, A.H.M., 2000. The survival motor neuron (SMN) protein: Effect of exon loss and mutation on protein localization. *Neurogenetics*. **3**, 7-16.
- Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L., Steitz, J.A., 1980. Are snRNPs involved in splicing? *Nature*. **283**, 220-224.
- Lewis, H., Perrett, A.J., Burley, G.A., Eperon, I.C., 2012. An RNA splicing enhancer that does not act by looping. *Angewandte Chemie - International Edition*. **51**, 9800-9803.
- Liu, H., Matsugami, A., Katahira, M., Uesugi, S., 2002. A dimeric RNA quadruplex architecture comprised of two G:G(:A):G:G(:A) hexads, G:G:G:G tetrads and UUUU loops. *Journal of Molecular Biology*. **322**, 955-970.
- Liu, M., Kumar, K.U., Pater, M.M., Pater, A., 1998. Identification and characterization of a JC virus pentanucleotide repeat element binding protein: Cellular nucleic acid binding protein. *Virus Research*. **58**, 73-82.
- Long, J.C. & Caceres, J.F., 2009. The SR protein family of splicing factors: Master regulators of gene expression. *Biochemical Journal*. **417**, 15-27.
- López-Bigas, N., Audit, B., Ouzounis, C., Parra, G., Guigó, R., 2005. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Letters*. **579**, 1900-1903.
- Lorson, C.L. & Androphy, E.J., 2000. An exonic enhancer is required for inclusion of an essential exon in the SMA-determining gene SMN. *Human Molecular Genetics*. **9**, 259-266.
- Lorson, C.L., Hahnen, E., Androphy, E.J., Wirth, B., 1999. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proceedings of the National Academy of Sciences of the United States of America*. **96**, 6307-6311.

BIBLIOGRAPHY

Lorson, C.L., Strasswimmer, J., Yao, J.-., Baleja, J.D., Hahnen, E., Wirth, B., Le, T., Burghes, A.H.M., Androphy, E.J., 1998. SMN oligomerization defect correlates with spinal muscular atrophy severity. *Nature Genetics*. **19**, 63-66.

Lu, M., Guo, Q., Kallenbach, N.R., 1992. Structure and stability of sodium and potassium complexes of dT4G4 and dT4G4T. *Biochemistry*[®]. **31**, 2455-2459.

Lunn, M.R. & Wang, C.H., 2008. Spinal muscular atrophy. *The Lancet*. **371**, 2120-2133.

Ma, X., Zhao, Y., Li, Y., Lu, H., He, Y., 2010. Relevance of Bcl-x expression in different types of endometrial tissues. *Journal of Experimental and Clinical Cancer Research*. **29**, .

Madhani, H.D. & Guthrie, C., 1992. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell*. **71**, 803-817.

Makarov, E.M., Owen, N., Bottrill, A., Makarova, O.V., 2012. Functional mammalian spliceosomal complex e contains SMN complex proteins in addition to U1 and U2 snRNPs. *Nucleic Acids Research*. **40**, 2639-2652.

Marcel, V., Tran, P.L.T., Sagne, C., Martel-Planche, G., Vaslin, L., Teulade-Fichou, M.-., Hall, J., Mergny, J.-., Hainaut, P., van Dyck, E., 2011. G-quadruplex structures in TP53 intron 3: Role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis*. **32**, 271-278.

Maroney, P.A., Romfo, C.M., Nilsen, T.W., 2000. Functional recognition of the 5' splice site by U4/U6.U5 tri-snRNP defines a novel ATP-dependent step in early spliceosome assembly. *Molecular Cell*. **6**, 317-328.

Mashima, T., Nishikawa, F., Kamatari, Y.O., Fujiwara, H., Saimura, M., Nagata, T., Kodaki, T., Nishikawa, S., Kuwata, K., Katahira, M., 2013. Anti-prion activity of an RNA aptamer and its structural basis. *Nucleic Acids Research*. **41**, 1355-1362.

Mattis, V.B., Ebert, A.D., Fosso, M.Y., Chang, C.-., Lorson, C.L., 2009. Delivery of a read-through inducing compound, TC007, lessens the severity of a spinal muscular atrophy animal model. *Human Molecular Genetics*. **18**, 3906-3913.

Mattis, V.B., Rai, R., Wang, J., Chang, C.-T., Coady, T., Lorson, C.L., 2006. Novel aminoglycosides increase SMN levels in spinal muscular atrophy fibroblasts. *Human Genetics*. **120**, 589-601.

Mayeda, A., Helfman, D.M., 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**, 2993-3001.

BIBLIOGRAPHY

- Mayeda, A. & Krainer, A.R., 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell*. **68**, 365-375.
- Mayes, A.E., Verdone, L., Legrain, P., Beggs, J.D., 1999. Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO Journal*. **18**, 4321-4331.
- McAndrew, P.E., Parsons, D.W., Simard, L.R., Rochette, C., Ray, P.N., Mendell, J.R., Prior, T.W., Burghes, A.H.M., 1997. Identification of proximal spinal muscular atrophy carriers and patients by analysis of SMN(T) and SMN(C) gene copy number. *American Journal of Human Genetics*. **60**, 1411-1422.
- McGlinchy, N.J., Valomon, A., Chesham, J.E., Maywood, E.S., Hastings, M.H., Ule, J., 2012. Regulation of alternative splicing by the circadian clock and food related cues. *Genome Biology*. **13**, .
- Merino, E.J., Wilkinson, K.A., Coughlan, J.L., Weeks, K.M., 2005. RNA structure analysis at single nucleotide resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE). *Journal of the American Chemical Society*. **127**, 4223-4231.
- Meyer, K., Marquis, J., Trüb, J., Nlend Nlend, R., Verp, S., Ruepp, M.-., Imboden, H., Barde, I., Trono, D., Schümperli, D., 2009. Rescue of a severe mouse model for spinal muscular atrophy by U7 snRNA-mediated splicing modulation. *Human Molecular Genetics*. **18**, 546-555.
- Michelotti, E.F., Tomonaga, T., Krutzsch, H., Levens, D., 1995. Cellular nucleic acid binding protein regulates the CT element of the human c-myc protooncogene. *Journal of Biological Chemistry*. **270**, 9494-9499.
- Millevoi, S., Moine, H., Vagner, S., 2012. G-quadruplexes in RNA biology. *Wiley Interdisciplinary Reviews: RNA*. **3**, 495-507.
- Miyajima, H., Miyaso, H., Okumura, M., Kurisu, J., Imaizumi, K., 2002. Identification of a cis-acting element for the regulation of SMN exon 7 splicing. *Journal of Biological Chemistry*. **277**, 23271-23277.
- Monani, U.R., 2005. Spinal muscular atrophy: A deficiency in a ubiquitous protein; a motor neuron-specific disease. *Neuron*. **48**, 885-896.
- Monani, U.R., Lorson, C.L., Parsons, D.W., Prior, T.W., Androphy, E.J., Burghes, A.H.M., McPherson, J.D., 1999. A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. *Human Molecular Genetics*. **8**, 1177-1183.
- Moore, M.J. & Sharp, P.A., 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature*. **365**, 364-368.

BIBLIOGRAPHY

- Morris, M.J., Negishi, Y., Pázsint, C., Schonhoff, J.D., Basu, S., 2010. An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES. *Journal of the American Chemical Society*. **132**, 17831-17839.
- Moser, H.E. & Dervan, P.B., 1987. Sequence-specific cleavage of double helical DNA by triple helix formation. *Science*. **238**, 645-650.
- Mount, S.M., 1982. A catalogue of splice junction sequences. *Nucleic Acids Research*. **10**, 459-472.
- Mourelatos, Z., Abel, L., Yong, J., Kataoka, N., Dreyfuss, G., 2001. SMN interacts with a novel family of hnRNP and spliceosomal proteins. *EMBO Journal*. **20**, 5443-5452.
- Mura, C., Cascio, D., Sawaya, M.R., Eisenberg, D.S., 2001. The crystal structure of a heptameric archaeal Sm protein: Implications for the eukaryotic snRNP core. *Proceedings of the National Academy of Sciences of the United States of America*. **98**, 5532-5537.
- Muraki, M., Ohkawara, B., Hosoya, T., Onogi, H., Koizumi, J., Koizumi, T., Sumi, K., Yomoda, J.-., Murray, M.V., Kimura, H., Furuichi, K., Shibuya, H., Krainer, A.R., Suzuki, M., Hagiwara, M., 2004. Manipulation of alternative splicing by a newly developed inhibitor of Clks. *Journal of Biological Chemistry*. **279**, 24246-24254.
- Murchie, A.I.H. & Lilley, D.M.J., 1994. Tetraplex folding of telomere sequences and the inclusion of adenine bases. *EMBO Journal*. **13**, 993-1001.
- Murphy, W.J., Watkins, K.P., Agabian, N., 1986. Identification of a novel Y branch structure as an intermediate in trypanosome mRNA processing: Evidence for trans splicing. *Cell*. **47**, 517-525.
- Nasim, F.-H., Hutchison, S., Cordeau, M., Chabot, B., 2002. High-affinity hnRNP A1 binding sites and duplex-forming inverted repeats have similar effects on 5' splice site selection in support of a common looping out and repression mechanism. *RNA*. **8**, 1078-1089.
- News-medical.net, 2013. *ISIS conducts phase 2 study of ISIS-SMNRx in infants with type I spinal muscular atrophy*. [online]. Available at: <http://www.news-medical.net/news/20131004/ISIS-conducts-phase-2-study-of-ISIS-SMNRx-in-infants-with-type-I-spinal-muscular-atrophy.aspx> [accessed November 05 2013].
- Okunola, H.L. & Krainer, A.R., 2009. Cooperative-binding and splicing-repressive properties of hnRNP A1. *Molecular and Cellular Biology*. **29**, 5620-5631.
- O'Mullane, L. & 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**, 7510-7520.

BIBLIOGRAPHY

- Osman, E.Y., Yen, P.-., Lorson, C.L., 2012. Bifunctional RNAs targeting the intronic splicing silencer N1 increase SMN levels and reduce disease severity in an animal model of spinal muscular atrophy. *Molecular Therapy*. **20**, 119-126.
- Otter, S., Grimmler, M., Neuenkirchen, N., Chari, A., Sickmann, A., Fischer, U., 2007. A comprehensive interaction map of the human survival of motor neuron (SMN) complex. *Journal of Biological Chemistry*. **282**, 5825-5833.
- Owen, N., Zhou, H., Malygin, A.A., Sangha, J., Smith, L.D., Muntoni, F., Eperon, I.C., 2011. Design principles for bifunctional targeted oligonucleotide enhancers of splicing. *Nucleic Acids Research*. 1-15.
- Pan, Q., Shai, O., Lee, L.J., Frey, B.J., Blencowe, B.J., 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics*. **40**, 1413-1415.
- Paronetto, M.P., Achsel, T., Massiello, A., Chalfant, C.E., Sette, C., 2007. The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x. *Journal of Cell Biology*. **176**, 929-939.
- Paronetto, M.P., Messina, V., Barchi, M., Geremia, R., Richard, S., Sette, C., 2011. Sam68 marks the transcriptionally active stages of spermatogenesis and modulates alternative splicing in male germ cells. *Nucleic Acids Research*. **39**, 4961-4974.
- Passini, M.A., Bu, J., Richards, A.M., Kinnecom, C., Sardi, S.P., Stanek, L.M., Hua, Y., Rigo, F., Matson, J., Hung, G., Kaye, E.M., Shihabuddin, L.S., Krainer, A.R., Bennett, C.F., Cheng, S.H., 2011. Antisense oligonucleotides delivered to the mouse CNS ameliorate symptoms of severe spinal muscular atrophy. *Science Translational Medicine*. **3**, .
- Passini, M.A., Bu, J., Roskelley, E.M., Richards, A.M., Sardi, S.P., O'Riordan, C.R., Klinger, K.W., Shihabuddin, L.S., Cheng, S.H., 2010. CNS-targeted gene therapy improves survival and motor function in a mouse model of spinal muscular atrophy. *Journal of Clinical Investigation*. **120**, 1253-1264.
- Patel, D.J., Phan, A.T., Kuryavyi, V., 2007. Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: Diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Research*. **35**, 7429-7455.
- Pearn, J., 1980. Classification of spinal muscular atrophies. *Lancet*. **1**, 919-922.
- Pedrotti, S., Bielli, P., Paronetto, M.P., Ciccocanti, F., Fimia, G.M., Stamm, S., Manley, J.L., Sette, C., 2010. The splicing regulator Sam68 binds to a novel exonic splicing silencer and functions in SMN2 alternative splicing in spinal muscular atrophy. *EMBO Journal*. **29**, 1235-1247.

BIBLIOGRAPHY

Perrett, A.J., Dickinson, R.L., Krpetić, Z., Brust, M., Lewis, H., Eperon, I.C., Burley, G.A., 2013. Conjugation of PEG and gold nanoparticles to increase the accessibility and valency of tethered RNA splicing enhancers. *Chemical Science*. **4**, 257-265.

Phan, A.T., Kuryavyi, V., Darnell, J.C., Serganov, A., Majumdar, A., Ilin, S., Raslin, T., Polonskaia, A., Chen, C., Clain, D., Darnell, R.B., Patel, D.J., 2011. Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. *Nature Structural and Molecular Biology*. **18**, 796-804.

Pilch, D.S., Plum, G.E., Breslauer, K.J., 1995. The thermodynamics of DNA structures that contain lesions or guanine tetrads. *Current Opinion in Structural Biology*. **5**, 334-342.

Pomeranz Krummel, D.A., Oubridge, C., Leung, A.K.W., Li, J., Nagai, K., 2009. Crystal structure of human spliceosomal U1 snRNP at 5.5 resolution. *Nature*. **458**, 475-480.

Porensky, P.N., Mitrpant, C., McGovern, V.L., Bevan, A.K., Foust, K.D., Kaspar, B.K., Wilton, S.D., Burghes, A.H., 2012. A single administration of morpholino antisense oligomer rescues spinal muscular atrophy in mouse. *Human Molecular Genetics*. **21**, 1625-1638.

Raal, F.J., Santos, R.D., Blom, D.J., Marais, A.D., Charng, M.-., Cromwell, W.C., Lachmann, R.H., Gaudet, D., Tan, J.L., Chasan-Taber, S., Tribble, D.L., Flaim, J.D., Crooke, S.T., 2010. Mipomersen, an apolipoprotein B synthesis inhibitor, for lowering of LDL cholesterol concentrations in patients with homozygous familial hypercholesterolaemia: a randomised, double-blind, placebo-controlled trial. *The Lancet*. **375**, 998-1006.

Rak, K., Lechner, B.D., Schneider, C., Drexler, H., Sendtner, M., Jablonka, S., 2009. Valproic acid blocks excitability in SMA type I mouse motor neurons. *Neurobiology of Disease*. **36**, 477-487.

Raker, V.A., Plessel, G., Lührmann, R., 1996. The snRNP core assembly pathway: Identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. *EMBO Journal*. **15**, 2256-2269.

Ramos, A., Hollingworth, D., Pastore, A., 2003. G-quartet-dependent recognition between the FMRP RGG box and RNA. *RNA*. **9**, 1198-1207.

Raponi, M., Buratti, E., Dassie, E., Upadhyaya, M., Baralle, D., 2009. Low U1 snRNP dependence at the NF1 exon 29 donor splice site. *FEBS Journal*. **276**, 2060-2073.

Reed, R., 1990. Protein composition of mammalian spliceosomes assembled in vitro. *Proceedings of the National Academy of Sciences of the United States of America*. **87**, 8031-8035.

BIBLIOGRAPHY

- Reed, R., 1989. The organization of 3' splice-site sequences in mammalian introns. *Genes and Development*. **3**, 2113-2123.
- Revil, T., Pelletier, J., Toutant, J., Cloutier, A., Chabot, B., 2009. Heterogeneous nuclear ribonucleoprotein K represses the production of pro-apoptotic Bcl-xS splice isoform. *Journal of Biological Chemistry*. **284**, 21458-21467.
- Rigo, F., Hua, Y., Chun, S.J., Prakash, T.P., Krainer, A.R., Bennett, C.F., 2012. Synthetic oligonucleotides recruit ILF2/3 to RNA transcripts to modulate splicing. *Nature Chemical Biology*. **8**, 555-561.
- Rigo, F., Hua, Y., Krainer, A.R., Frank Bennett, C., 2012. Antisense-based therapy for the treatment of spinal muscular atrophy. *Journal of Cell Biology*. **199**, 21-25.
- Rimessi, P., Sabatelli, P., Fabris, M., Braghetta, P., Bassi, E., Spitali, P., Vattei, G., Tomelleri, G., Mari, L., Perrone, D., Medici, A., Neri, M., Bovolenta, M., Martoni, E., Maraldi, N.M., Gualandi, F., Merlini, L., Ballestri, M., Tondelli, L., Sparnacci, K., Bonaldo, P., Caputo, A., Laus, M., Ferlini, A., 2009. Cationic PMMA nanoparticles bind and deliver antisense oligoribonucleotides allowing restoration of dystrophin expression in the mdx mouse. *Molecular Therapy*. **17**, 820-827.
- Rinke, J., Appel, B., Digweed, M., Luhrmann, R., 1985. Localization of a base-paired interaction between small nuclear RNAs U4 and U6 in intact U4/U6 ribonucleoprotein particles by psoralen cross-linking. *Journal of Molecular Biology*. **185**, 721-731.
- Robberson, B.L., Cote, G.J., Berget, S.M., 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Molecular and Cellular Biology*. **10**, 84-94.
- Roberts, G.C., Gooding, C., Mak, H.Y., Proudfoot, N.J., Smith, C.W.J., 1998. Co-transcriptional commitment to alternative splice site selection. *Nucleic Acids Research*. **26**, 5568-5572.
- Roca, X., Akerman, M., Gaus, H., Berdeja, A., Bennett, C.F., Krainer, A.R., 2012. Widespread recognition of 5' splice sites by noncanonical base-pairing to U1 snRNA involving bulged nucleotides. *Genes and Development*. **26**, 1098-1109.
- Roca, X., Krainer, A.R., Eperon, I.C., 2013. Pick one, but be quick: 5' splice sites and the problems of too many choices. *Genes and Development*. **27**, 129-144.
- Roecigno, R.F., Weiner, M., Garcia-Blanco, M.A., 1993. A mutational analysis of the polypyrimidine tract of introns: Effects of sequence differences in pyrimidine tracts on splicing. *Journal of Biological Chemistry*. **268**, 11222-11229.
- Rogers, J. & Wall, R., 1980. A mechanism for RNA splicing. *Proceedings of the National Academy of Sciences of the United States of America*. **77**, 1877-1879.

BIBLIOGRAPHY

- Rudner, D.Z., Breger, K.S., Rio, D.C., 1998. Molecular genetic analysis of the heterodimeric splicing factor U2AF: The RS domain on either the large or small *Drosophila* subunit is dispensable in vivo. *Genes and Development*. **12**, 1010-1021.
- Ruskin, B., Krainer, A.R., Maniatis, T., Green, M.R., 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell*. **38**, 317-331.
- Saccà, B., Lacroix, L., Mergny, J.-., 2005. The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides. *Nucleic Acids Research*. **33**, 1182-1192.
- Sakharkar, M.K., Perumal, B.S., Sakharkar, K.R., Kanguane, P., 2005. An analysis on gene architecture in human and mouse genomes. *In Silico Biology*. **5**, 347-365.
- Sakla, M.S. & Lorson, C.L., 2008. Induction of full-length survival motor neuron by polyphenol botanical compounds. *Human Genetics*. **122**, 635-643.
- Sanford, J.R., Wang, X., Mort, M., Vanduyne, N., Cooper, D.N., Mooney, S.D., Edenberg, H.J., Liu, Y., 2009. Splicing factor SFRS1 recognizes a functionally diverse landscape of RNA transcripts. *Genome Research*. **19**, 381-394.
- Saulière, J., Sureau, A., Expert-Bezançon, A., Marie, J., 2006. The polypyrimidine tract binding protein (PTB) represses splicing of exon 6B from the β -tropomyosin pre-mRNA by directly interfering with the binding of the U2AF65 subunit. *Molecular and Cellular Biology*. **26**, 8755-8769.
- Sazani, P., Ness, K.P.V., Weller, D.L., Poage, D.W., Palyada, K., Shrewsbury, S.B., 2011. Repeat-dose toxicology evaluation in cynomolgus monkeys of AVI-4658, a phosphorodiamidate morpholino oligomer (PMO) drug for the treatment of duchenne muscular dystrophy. *International Journal of Toxicology*. **30**, 313-321.
- Schad, E., Tompa, P., Hegyi, H., 2011. The relationship between proteome size, structural disorder and organism complexity. *Genome Biology*. **12**, .
- Schneider, C., Will, C.L., Makarova, O.V., Makarov, E.M., Lührmann, R., 2002. Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions. *Molecular and Cellular Biology*. **22**, 3219-3229.
- Schwartz, S., Meshorer, E., Ast, G., 2009. Chromatin organization marks exon-intron structure. *Nature Structural and Molecular Biology*. **16**, 990-995.
- Seela, F., Chen, Y., Mittelbach, C., 1998. 7-Deazapurine Oligodeoxyribonucleotides: The Effects of 7-Deaza-8-methylguanine on DNA Structure and Stability. *Helvetica Chimica Acta*. **81**, 570-583.

BIBLIOGRAPHY

- Seiwert, S.D. & Steitz, J.A., 1993. Uncoupling two functions of the U1 small nuclear ribonucleoprotein particle during in vitro splicing. *Molecular and Cellular Biology*. **13**, 3135-3145.
- Selenko, P., Gregorovic, G., Sprangers, R., Stier, G., Rhani, Z., Krämer, A., Sattler, M., 2003. Structural basis for the molecular recognition between human splicing factors U2AF65 and SF1/mBBP. *Molecular Cell*. **11**, 965-976.
- Sen, D. & Gilbert, W., 1992. Novel DNA superstructures formed by telomere-like oligomers. *Biochemistry*. **31**, 65-70.
- Shapiro, M.B. & Senapathy, P., 1987. RNA splice junctions of different classes of eukaryotes: Sequence statistics and functional implications in gene expression. *Nucleic Acids Research*. **15**, 7155-7174.
- Sharma, S., Falick, A.M., Black, D.L., 2005. Polypyrimidine tract binding protein blocks the 5' splice site-dependent assembly of U2AF and the prespliceosomal e complex. *Molecular Cell*. **19**, 485-496.
- Sharma, S., Maris, C., Allain, F.-., Black, D., 2011. U1 snRNA Directly Interacts with Polypyrimidine Tract-Binding Protein during Splicing Repression. *Molecular Cell*. **41**, 579-588.
- Sharp, P.A., 2005. The discovery of split genes and RNA splicing. *Trends in Biochemical Sciences*. **30**, 279-281.
- Shen, H. & Green, M.R., 2006. RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. *Genes and Development*. **20**, 1755-1765.
- Shen, H. & Green, M.R., 2004. A pathway of sequential Arginine-Serine-Rich Domain-Splicing signal interactions during mammalian spliceosome assembly. *Molecular Cell*. **16**, 363-373.
- Shen, H., Kan, J.L.C., Green, M.R., 2004. Arginine-Serine-Rich Domains Bound at Splicing Enhancers Contact the Branchpoint to Promote Prespliceosome Assembly. *Molecular Cell*. **13**, 367-376.
- Shin, C. & Manley, J.L., 2004. Cell signalling and the control of pre-mRNA splicing. *Nature Reviews Molecular Cell Biology*. **5**, 727-738.
- Silva, J.C., Gorenstein, M.V., Li, G.-., Vissers, J.P.C., Geromanos, S.J., 2006. Absolute quantification of proteins by LCMSE: A virtue of parallel MS acquisition. *Molecular and Cellular Proteomics*. **5**, 144-156.
- Simonsson, T., 2001. G-quadruplex DNA structures - Variations on a theme. *Biological Chemistry*. **382**, 621-628.

BIBLIOGRAPHY

- Singh, N.K., Singh, N.N., Androphy, E.J., Singh, R.N., 2006. Splicing of a critical exon of human Survival Motor Neuron is regulated by a unique silencer element located in the last intron. *Molecular and Cellular Biology*. **26**, 1333-1346.
- Singh, N.N., Androphy, E.J., Singh, R.N., 2004. In vivo selection reveals combinatorial controls that define a critical exon in the spinal muscular atrophy genes. *RNA*. **10**, 1291-1305.
- Singh, N.N., Shishimorova, M., Lu, C.C., Gangwani, L., Singh, R.N., 2009. A short antisense oligonucleotide masking a unique intronic motif prevents skipping of a critical exon in spinal muscular atrophy. *RNA Biology*. **6**, 130-139.
- Singh, N.N., Singh, R.N., Androphy, E.J., 2007. Modulating role of RNA structure in alternative splicing of a critical exon in the spinal muscular atrophy genes. *Nucleic Acids Research*. **35**, 371-389.
- Singh, R., Valcarcel, J., Green, M.R., 1995. Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science*. **268**, 1173-1176.
- Skordis, L.A., Dunckley, M.G., Burglen, L., Campbell, L., Talbot, K., Patel, S., Melki, J., Davies, K.E., Dubowitz, V., Muntoni, F., 2001. Characterisation of novel point mutations in the survival motor neuron gene SMN, in three patients with SMA. *Human Genetics*. **108**, 356-357.
- Skordis, L.A., Dunckley, M.G., Yue, B., Eperon, I.C., Muntoni, F., 2003. Bifunctional antisense oligonucleotides provide a trans-acting splicing enhancer that stimulates SMN2 gene expression in patient fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America*. **100**, 4114-4119.
- Skotheim, R.I. & Nees, M., 2007. Alternative splicing in cancer: Noise, functional, or systematic? *International Journal of Biochemistry and Cell Biology*. **39**, 1432-1449.
- Smith, L.D., 2012. *An investigation into the mechanism of action of the SMN2 bifunctional oligonucleotide and its application to reverse the RON Δ 165 pro-metastatic splicing event*, <http://hdl.handle.net/2381/10826>. Ph. D. University of Leicester.
- Smith, L.D., Dickinson, R.L., Cousins, A., Malygin, A.A., Lucas, C.M., Perrett, A.J., Searle, M.S., Burley, G.A., Eperon, I.C., manuscript submitted. Targeted bifunctional oligonucleotide enhancers of SMN2 exon 7 splicing recruit U2AF65 via the annealing domain and U2 snRNP via the quadruplex-forming ESE domain.
- Smith, C.W.J., Porro, E.B., Patton, J.G., Nadal-Ginard, B., 1989. Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. *Nature*. **342**, 243-247.
- Solnick, D., 1985. Alternative splicing caused by RNA secondary structure. *Cell*. **43**, 667-676.

BIBLIOGRAPHY

- Sorek, R. & Ast, G., 2003. Intronic sequences flanking alternatively spliced exons are conserved between human and mouse. *Genome Research*. **13**, 1631-1637.
- Soret, J., Bakkour, N., Maire, S., Durand, S., Zekri, L., Gabut, M., Fic, W., Divita, G., Rivalle, C., Dauzonne, D., Nguyen, C.H., Jeanteur, P., Tazi, J., 2005. Selective modification of alternative splicing by indole derivatives that target serine-arginine-rich protein splicing factors. *Proceedings of the National Academy of Sciences of the United States of America*. **102**, 8764-8769.
- Spellman, R. & Smith, C.W.J., 2006. Novel modes of splicing repression by PTB. *Trends in Biochemical Sciences*. **31**, 73-76.
- Spena, S., Tenchini, M.L., Buratti, E., 2006. Cryptic splice site usage in exon 7 of the human fibrinogen B β -chain gene is regulated by a naturally silent SF2/ASF binding site within this exon. *RNA*. **12**, 948-958.
- Spies, N., Nielsen, C.B., Padgett, R.A., Burge, C.B., 2009. Biased Chromatin Signatures around Polyadenylation Sites and Exons. *Molecular Cell*. **36**, 245-254.
- Staknis, D. & Reed, R., 1994. SR proteins promote the first specific recognition of pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. *Molecular and Cellular Biology*. **14**, 7670-7682.
- Staley, J.P. & Guthrie, C., 1998. Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell*. **92**, 315-326.
- Stamm, S., 2002. Signals and their transduction pathways regulating alternative splicing: A new dimension of the human genome. *Human Molecular Genetics*. **11**, 2409-2416.
- Stark, H., Dube, P., Lührmann, R., Kastner, B., 2001. Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature*. **409**, 539-542.
- Sterner, D.A., Carlo, T., Berget, S.M., 1996. Architectural limits on split genes. *Proceedings of the National Academy of Sciences of the United States of America*. **93**, 15081-15085.
- Subramanian, M., Rage, F., Tabet, R., Flatter, E., Mandel, J.-., Moine, H., 2011. G-quadruplex RNA structure as a signal for neurite mRNA targeting. *EMBO Reports*. **12**, 697-704.
- Sugnet, C.W., Kent, W.J., Ares Jr., M., Haussler, D., 2004. Transcriptome and genome conservation of alternative splicing events in humans and mice. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*. 66-77.
- Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki, Y., Nakamura, Y., Suyama, A., Sugano, S.,

BIBLIOGRAPHY

2001. Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites. *EMBO Reports*. **2**, 388-393.
- Tange, T.Ø, Damgaard, C.K., Guth, S., Valcárcel, J., Kjems, J., 2001. The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. *EMBO Journal*. **20**, 5748-5758.
- Tarn, W.-. & Steitz, J.A., 1994. SR proteins can compensate for the loss of U1 snRNP functions in vitro. *Genes and Development*. **8**, 2704-2717.
- Tavanez, J., Madl, T., Kooshapur, H., Sattler, M., Valcárcel, J., 2012. HnRNP A1 Proofreads 3' Splice Site Recognition by U2AF. *Molecular Cell*. **45**, 314-329.
- Taylor, J.E., Thomas, N.H., Lewis, C.M., Abbs, S.J., Rodrigues, N.R., Davies, K.E., Mathew, C.G., 1998. Correlation of SMNt and SMNc gene copy number with age of onset and survival in spinal muscular atrophy. *European Journal of Human Genetics*. **6**, 467-474.
- Tian, M. & Maniatis, T., 1994. A splicing enhancer exhibits both constitutive and regulated activities. *Genes and Development*. **8**, 1703-1712.
- Tian, M. & Maniatis, T., 1993. A splicing enhancer complex controls alternative splicing of doublesex pre-mRNA. *Cell*. **74**, 105-114.
- Tian, M. & Maniatis, T., 1992. Positive control of pre-mRNA splicing in vitro. *Science*. **256**, 237-240.
- Todd, A.K., Johnston, M., Neidle, S., 2005. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Research*. **33**, 2901-2907.
- Treisman, R., Orkin, S.H., Maniatis, T., 1983. Structural and functional defects in beta-thalassemia. *Progress in Clinical and Biological Research*. **134**, 99-121.
- Turunen, J.J., Niemelä, E.H., Verma, B., Frilander, M.J., 2013. The significant other: Splicing by the minor spliceosome. *Wiley Interdisciplinary Reviews: RNA*. **4**, 61-76.
- Uesugi, S., Liu, H., Kugimiya, A., Matsugami, A., Katahira, M., 2003. RNA and DNA, which contain two GGAGG segments connected with UUUU or TTTT sequences, form entirely different quadruplex structures. *Nucleic Acids Research. Supplement (2001)*. 51-52.
- Valcárcel, J., Gaur, R.K., Singh, R., Green, M.R., 1996. Interaction of U2AF65 RS region with pre-mRNA of branch point and promotion base pairing with U2 snRNA. *Science*. **273**, 1706-1709.

BIBLIOGRAPHY

- Valori, C.F., Ning, K., Wyles, M., Mead, R.J., Grierson, A.J., Shaw, P.J., Azzouz, M., 2010. Systemic delivery of scAAV9 expressing SMN prolongs survival in a model of spinal muscular atrophy. *Science Translational Medicine*. **2**, .
- Vaughn, J.P., Creacy, S.D., Routh, E.D., Joyner-Butt, C., Jenkins, G.S., Pauli, S., Nagamine, Y., Akman, S.A., 2005. The DEXH protein product of the DHX36 gene is the major source of tetramolecular quadruplex G4-DNA resolving activity in HeLa cell lysates. *Journal of Biological Chemistry*. **280**, 38117-38120.
- Venkatesh, S., Wower, J., Byrne, M.E., 2009. Nucleic acid therapeutic carriers with on-demand triggered release. *Bioconjugate Chemistry*. **20**, 1773-1782.
- Vinogradov, S.V., Batrakova, E.V., Kabanov, A.V., 2004. Nanogels for Oligonucleotide Delivery to the Brain. *Bioconjugate Chemistry*. **15**, 50-60.
- Wahl, M.C., Will, C.L., Lührmann, R., 2009. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell*. **136**, 701-718.
- Wang, E.T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., Burge, C.B., 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature*. **456**, 470-476.
- Wang, J. & Manley, J.L., 1995. Overexpression of the SR proteins ASF/SF2 and SC35 influences alternative splicing in vivo in diverse ways. *RNA*. **1**, 335-346.
- Wang, Z. & Burge, C.B., 2008. Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *RNA*. **14**, 802-813.
- Wang, Z., Hoffmann, H.M., Grabowski, P.J., 1995. Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA*. **1**, 21-35.
- Wang, Z., Xiao, X., Van Nostrand, E., Burge, C.B., 2006. General and Specific Functions of Exonic Splicing Silencers in Splicing Control. *Molecular Cell*. **23**, 61-70.
- Wassarman, D.A. & Steitz, J.A., 1992. Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. *Science*. **257**, 1918-1925.
- Wieland, M. & Hartig, J.S., 2007. RNA Quadruplex-Based Modulation of Gene Expression. *Chemistry and Biology*. **14**, 757-763.
- Wieringa, B., Meyer, F., Reiser, J., Weissmann, C., 1983. Unusual splice sites revealed by mutagenic inactivation of an authentic splice site of the rabbit β -globin gene. *Nature*. **301**, 38-43.
- Wilkinson, K.A., Merino, E.J., Weeks, K.M., 2006. Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): Quantitative RNA structure analysis at single nucleotide resolution. *Nature Protocols*. **1**, 1610-1616.

BIBLIOGRAPHY

- Will, C.L. & Lührmann, R., 2005. Splicing of a rare class of introns by the U12-dependent spliceosome. *Biological Chemistry*. **386**, 713-724.
- Williams, J.H., Schray, R.C., Patterson, C.A., Ayitey, S.O., Tallent, M.K., Lutz, G.J., 2009. Oligonucleotide-mediated survival of motor neuron protein expression in CNS improves phenotype in a mouse model of spinal muscular atrophy. *Journal of Neuroscience*. **29**, 7633-7638.
- Wirth, B., 2000. An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Human Mutation*. **15**, 228-237.
- Wojtowicz, W.M., Wu, W., Andre, I., Qian, B., Baker, D., Zipursky, S.L., 2007. A Vast Repertoire of Dscam Binding Specificities Arises from Modular Interactions of Variable Ig Domains. *Cell*. **130**, 1134-1145.
- Wu, J.Y. & Maniatis, T., 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*. **75**, 1061-1070.
- Wu, S., Romfo, C.M., Nilsen, T.W., Green, M.R., 1999. Functional recognition 3' splice site AG by the splicing factor U2AF35. *Nature*. **402**, 832-835.
- Yang, X., Bani, M.-., Lu, S.-., Rowan, S., Ben-David, Y., Chabot, B., 1994. The A1 and A1(B) proteins of heterogeneous nuclear ribonucleoproteins modulate 5' splice site selection in vivo. *Proceedings of the National Academy of Sciences of the United States of America*. **91**, 6924-6928.
- Yang, X., Koh, C.G., Liu, S., Pan, X., Santhanam, R., Yu, B., Peng, Y., Pang, J., Golan, S., Talmon, Y., Jin, Y., Muthusamy, N., Byrd, J.C., Chan, K.K., Lee, L.J., Marcucci, G., Lee, R.J., 2009. Transferrin receptor-targeted lipid nanoparticles for delivery of an antisense oligodeoxyribonucleotide against bcl-2. *Molecular Pharmaceutics*. **6**, 221-230.
- Yeo, G., Holste, D., Kreiman, G., Burge, C.B., 2004. Variation in alternative splicing across human tissues. *Genome Biology*. **5**, .
- Yuo, C.-., Lin, H.-., Chang, Y.-., Yang, W.-., Chang, J.-., 2008. 5-(N-ethyl-N-isopropyl)-amiloride enhances SMN2 exon 7 inclusion and protein expression in spinal muscular atrophy cells. *Annals of Neurology*. **63**, 26-34.
- Zahler, A.M., Damgaard, C.K., Kjems, J., Caputi, M., 2004. SC35 and Heterogeneous Nuclear Ribonucleoprotein A/B Proteins Bind to a Juxtaposed Exonic Splicing Enhancer/Exonic Splicing Silencer Element to Regulate HIV-1 tat Exon 2 Splicing. *Journal of Biological Chemistry*. **279**, 10077-10084.
- Zahler, A.M. & 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 of the United States of America*. **92**, 2642-2646.

BIBLIOGRAPHY

- Zamore, P.D. & Green, M.R., 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proceedings of the National Academy of Sciences of the United States of America*. **86**, 9243-9247.
- Zarudnaya, M.I., Kolomiets, I.M., Potyahaylo, A.L., Hovorun, D.M., 2003. Downstream elements of mammalian pre-mRNA polyadenylation signals: Primary, secondary and higher-order structures. *Nucleic Acids Research*. **31**, 1375-1386.
- Zhang, M.Q., 1998. Statistical features of human exons and their flanking regions. *Human Molecular Genetics*. **7**, 919-932.
- Zhang, Q.-., Manche, L., Xu, R.-., Krainer, A.R., 2006. hnRNP A1 associates with telomere ends and stimulates telomerase activity. *RNA*. **12**, 1116-1128.
- Zhang, T., Zhang, H., Wang, Y., McGown, L.B., 2012. Capture and identification of proteins that bind to a GGA-rich sequence from the ERBB2 gene promoter region. *Analytical and Bioanalytical Chemistry*. **404**, 1867-1876.
- Zhang, X.H.-. & Chasin, L.A., 2004. Computational definition of sequence motifs governing constitutive exon splicing. *Genes and Development*. **18**, 1241-1250.
- Zhong, X.-., Wang, P., Han, J., Rosenfeld, M.G., Fu, X.-., 2009. SR Proteins in Vertical Integration of Gene Expression from Transcription to RNA Processing to Translation. *Molecular Cell*. **35**, 1-10.
- Zhou, H., Janghra, N., Mitrpant, C., Dickinson, R.L., Anthony, K., Price, L., Eperon, I.C., Wilton, S.D., Morgan, J., Muntoni, F., 2013. A Novel Morpholino Oligomer Targeting ISS-N1 Improves Rescue of Severe Spinal Muscular Atrophy Transgenic Mice. *Human Gene Therapy*.
- Zhu, J. & Krainer, A.R., 2000. Pre-mRNA splicing in the absence of an SR protein RS domain. *Genes and Development*. **14**, 3166-3178.
- Zhu, J., Mayeda, A., Krainer, A.R., 2001. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Molecular Cell*. **8**, 1351-1361.
- Zieve, G.W. & Khusial, P.R., 2003. The anti-Sm immune response in autoimmunity and cell biology. *Autoimmunity Reviews*. **2**, 235-240.
- Zuo, P. & Maniatis, T., 1996. The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. *Genes and Development*. **10**, 1356-1368.