

**MODIFICATION OF SPLICING PATTERNS USING
ARTIFICIAL ENHANCER SEQUENCES WITHIN
TRIPARTITE STRUCTURE**

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Andrew Jonathan Perrett MChem

Department of Chemistry

University of Leicester

2012

Abstract

Conjugation of linkers and gold nanoparticles to increase the activity of tethered RNA splicing enhancers

Perrett, A. J.; Dickinson, R. L.; Krpetić, Ž.; Brust, M.; Lewis, H.; Eperon, I. C.; Burley, G. A.

Genes *SMN1* and *SMN2* encode SMN protein which is essential for motor neurone survival. In Spinal Muscular Atrophy patients, absence of functional *SMN1* means functional SMN protein is encoded by *SMN2* alone. *SMN2* splicing products do not usually include exon 7, which is required for functional SMN production. Therefore, methods for increasing exon 7 inclusion levels of *SMN2* spliced products were investigated. 2'OMe RNA annealer and enhancer sequences were synthesised based on Targeted Oligonucleotide Enhancers of Splicing (TOES) with the addition of HEG linkers and bioconjugation intermediates. Cu-catalysed click chemistry was used to conjugate these sequences together to form tripartite sequences. Tripartite sequences RNA3-RNA11 were synthesised along with phosphothioate (pS) versions of RNA6 and RNA7. Adding a bioconjugation group between enhancer and annealer strands increased exon 7 inclusion above basal levels in splicing experiments. pS tripartite sequences were less effective than non-pS, versions, possibly due pS backbones affecting the strength of binding between annealers and mRNA. RNA6 and RNA7, containing 1 and 2 HEG linkers respectively, achieved exon 7 inclusion in 40% and 35% of *SMN2* spliced products, respectively. GGA-O only achieved 26% exon 7 inclusion, with basal levels at 11%. Less flexible abasic linkers equivalent in length to one HEG unit, gave lower levels of exon 7 inclusion than those achieved with HEG, suggesting the flexibility of HEG linkers is important for enhancing exon inclusion. Multivalent systems were constructed consisting of annealer and enhancer sequences polyconjugated to gold nanoparticles (GNPs). Splicing experiments showed exon 7 inclusion levels were dependent on GNP size, surface coating and directionality of strand attachment. GNP16 with an enhancer coating and 5nm diameter achieved exon 7 inclusion in 80% of spliced products, with GGA-O achieving 67% and basal levels at 44%. Addition of annealers to GNP16, significantly reduced exon 7 inclusion.

Acknowledgements

I would like to thank my supervisors Glenn and Ian for the opportunity that they gave me to work on this project. I also want to thank Paul Cullis for his support and help during the last few months of my PhD when Glenn was up in Stratheclyde. I want to thank all my friends in the second floor organic lab for their help and support throughout my PhD especially Helen, Vanessa and Rachel. I also want to thank Mark and Lindsay who helped me get up to speed with the biology and showed me the ropes. Most of all I would like to my lovely wife Laura for help and her support throughout the duration of my PhD and for providing for me while I was writing my thesis even when she was not feeling very well. Lastly, I would like to thank everybody I consulted and helped me in any way throughout these years especially Alex Cousins for collaborations on the quadruplex work and Željka Krpetić for her help with GNP characterisation.

Contents

Abstract	ii
Acknowledgements	iii
Dissemination of Results from this PhD Project.....	xiii
List of Tables	xiv
List of Figures	xvii
List of Schemes	xxii
1 - Introduction.....	1
1.1 - Alternative Splicing Mechanism	1
1.1.1 – H Complex Formation.....	4
1.1.2 – A Complex Formation.....	9
1.1.3 – Splicing Complexes B, B* and C.....	10
1.1.4 – Enhancers	13
1.1.4.1 – 5'SS Enhancement by SR Proteins	14
1.1.4.2 – 3'SS Enhancement by SR Proteins	15
1.1.5 – Silencers	16
1.1.5.1 – SS Inhibition by hnRNP	17
1.1.6 – Proximity Effect.....	17
1.1.7 – 3'SS Selection	19
1.1.8 – Exon Definition to Intron Definition.....	19
1.2 – Alternative Splicing and Disease	21

1.3 – Spinal Muscular Atrophy (SMA).....	22
1.4 – TOES.....	27
1.5 – Aims	30
1.6 – Hypothesis.....	31
Chapter 2 – Preparation of tripartite oligonucleotides	32
2 – Introduction.....	33
2.1 – Aims	33
2.2 – TOES Optimisations	34
2.2.1 – TOES Annealer Optimisations.....	34
2.2.2 – TOES Enhancer Optimisations	37
2.2.3 – Tripartite Sequence	39
2.2.4 – Bioconjugation Chemistry for Tripartite Sequences	41
2.2.4.1 – Amide Coupling.....	42
2.2.4.2 – Thiol-Maleimide Coupling.....	42
2.2.4.3 – Copper Catalysed Huisgen [3+2] Cycloaddition (Click Chemistry).....	42
2.3 – Strategy	45
2.4 – Results.....	48
2.4.1 – Linker Synthesis.....	48
2.4.2 – Preparation of Reactive Groups for Bioconjugation	51
2.4.3 – Synthesis of Enhancer and Annealer Strands for Bioconjugation	52
2.4.3.1 – 2'OMe RNA Solid Phase Synthesis.....	52

2.4.3.2 – Modification of the 3' Amino Modified Enhancers by NHS Azide Conjugation (Compound 11)	63
2.4.4 – Click Chemistry Reactions.....	65
2.4.4.1 – Trial Click Chemistry.....	65
2.4.5 – All 2'OMe RNA Click Chemistry Reactions.....	68
2.4.6 - 2'OMe RNA Click Chemistry Scale Up.....	73
2.4.7 – Phosphothioate Synthesis.....	75
2.4.8 – pS 2'OMe RNA Solid Phase Synthesis.....	76
2.4.9 – pS Trial Click Chemistry	79
2.4.10 – pS RNA6 & pS RNA7 Click Chemistry Scale Up.....	83
2.4.11 – Solid Phase Synthesis of GGA Chemistry Strands	84
2.4.12 – Quadruplex Study	87
2.4.13 – Discussion	94
2.4.14 – Conclusion	99
2.5 – Experimental	99
2.5.1 – General Procedures	100
2.5.1.1 – Standard DNA/RNA Synthesis	100
2.5.1.2 – HPLC Protocol.....	100
2.5.1.3 – Gel Imaging.....	101
2.5.1.4 – Desalting	101
2.5.1.5 – Click Chemistry Scale Up Protocol	101
2.5.2 – Synthesis of 2'OMe RNA Conjugates	102

2.5.2.1 – Preparation of 1,1-bis(4-methoxyphenyl)-1-phenyl-2,5,8,11,14,17-hexaoxonadecan-19-ol (12).....	102
2.5.2.2 – Preparation of Bioconjugation Reactants for Click Chemistry	103
2.5.2.3 – Preparation of 2-cyanoethyl N, N diisopropyl hex-5-yn-1-yl phosphoramidite (10) ¹⁵⁹	108
2.5.2.4 – Solid Phase Synthesis on a 0.2μmol for RNA1 and RNA2 Series.....	108
2.5.2.5 – Solid Phase Synthesis on a 1μmol for RNA1 and RNA2 Series	110
2.5.2.6 – Preparation of 2,5-dioxopyrrolidin-1-yl 4-azidobutanoate (11) ¹⁴⁸	112
2.5.2.7 – Preparation of RNA1_B from RNA1_A and Compound 11	113
2.5.2.8 – Preparation of RNA1_D from RNA1_C and Compound 11	113
2.5.2.9 – Preparation of RNA1_F from RNA1_E and Compound 11	114
2.5.3 – Click Chemistry of 2'OMe RNA Sequences	114
2.5.3.1 – Preparation of RNA6 Tripartite Structure.....	116
2.5.3.2 – Preparation of RNA11 Tripartite Structures	116
2.5.3.3 – All 2'OMe RNA Click Chemistry Reactions.....	117
2.5.4 – Click Chemistry Scale Up Reactions	118
2.5.4.1 – Trial Purification Condition of Scale Up Click.....	118
2.5.4.2 – RNA3 Scale Up.....	121
2.5.4.3 – RNA4 Scale Up.....	121
2.5.4.4 – RNA5 Scale Up.....	122
2.5.4.5 – RNA6 Scale Up.....	122
2.5.4.6 – RNA7 Scale Up.....	123

2.5.4.7 – RNA8 Scale Up.....	123
2.5.4.8 – RNA9 Scale Up.....	123
2.5.4.9 – RNA10 Scale Up.....	124
2.5.4.10 – RNA11 Scale Up.....	124
2.5.5 – 2'OMe Phosphorothioate (pS) RNA	125
2.5.5.1 – Synthesis of pS DNA 3	127
2.5.5.2 – Preparation of pS 2'OMe RNA Strands	127
2.5.5.3 – Preparation of 2'OMe pS RNA Azides.....	128
2.5.6 – All pS Click Chemistry Reactions	129
2.5.7 – pS Scale Up Click	132
2.5.7.1 – pS RNA6 Scale Up.....	132
2.5.7.2 – pS RNA7 Scale Up.....	133
2.5.8 – Synthesis of Abasic Strands	133
2.5.8.1 – Preparation of (2R,3S)-pentane-1,2,3,5-tetraol (14)	135
2.5.8.2 – Preparation of (2R,3S)-2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)tetrahydrofuran-3-ol (16)	136
2.5.8.3 – Preparation of 2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)tetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (17)	137
2.5.8.4 – Preparation of 2'OMe RNA Abasic Strands	138
2.5.8.5 – Preparation of RNA1_H Conjugate	139
2.5.9 – RNA15 Scale Up.....	139
2.5.10 – Solid Phase Synthesis of RNA16.....	141

Chapter 3 – Tripartite oligonucleotide splicing.....	144
3 – Introduction.....	145
3.1 – Aims	145
3.2 – Blocking Regulatory Sequences to Modify Splicing Outcomes	146
3.3 – Introducing Positive Factors	148
3.4 – Hypothesis.....	150
3.5 – Results.....	151
3.5.1 – <i>SMN2</i> Splicing	151
3.5.1.1 – 2'OMe RNA <i>SMN2</i> Splicing.....	151
3.5.1.2 – Increasing Protein Binding.....	155
3.5.1.3 – Constrain Freedom of Movement of the Linker and its Impact on Splicing.....	158
3.5.1.4 – GGA Modifications and Chemistry	162
3.6 – Discussion	164
3.7 – Conclusion	169
3.8 – Experimental	169
3.8.1 – Sequences.....	169
3.8.1.1 – PCR Sequence	169
3.8.1.2 – Pre-mRNA Sequence	170
3.8.2 – Standard Protocols	170
3.8.2.1 – ³² P-radiolabelled (Hot) Transcription Protocol.....	170
3.8.2.2 – Standard <i>in vitro</i> Splicing Reagents	171
3.8.2.3 – Standard <i>in vitro</i> Splicing Protocol	172

3.8.3 - PCR.....	173
3.8.4 – <i>In vitro</i> 2'OMe RNA Splicing.....	175
3.8.4.1 – <i>SMN2</i> Transcription	175
3.8.4.2 – <i>SMN2</i> Splicing Varying the Amount of NE	175
3.8.4.3 – <i>SMN2</i> Transcription	176
3.8.4.4 – 100nM Triplicate <i>SMN2</i> Splicing	176
3.8.4.5 – <i>SMN2</i> Transcription	176
3.8.4.6 – 50nM Triplicate <i>SMN2</i> Splicing	177
3.8.5 – <i>In vitro</i> 2'OMe pS RNA Splicing	177
3.8.5.1 – <i>SMN2</i> Transcription	177
3.8.5.2 – pS Splicing	178
3.8.6 – <i>In vitro</i> 2'OMe Abasic Splicing.....	178
3.8.6.1 – <i>SMN2</i> Transcription	178
3.8.6.2 – 70°C Annealing	179
3.8.6.3 – 30°C Annealing	179
3.8.7 – <i>In vitro</i> RNA Splicing	180
3.8.7.1 – <i>SMN2</i> Transcription	180
3.8.7.2 – RNA16 Splicing.....	180
Chapter 4 – Rectifying splicing defeats using polyconjugated GNP.....	182
4 – Introduction.....	183
4.1 – Aims	185
4.2 – Hypothesis.....	186

4.3 – Strategy	186
4.4 – Results	189
4.4.1 – Preparation of Strands for Polyconjugation	189
4.4.2 – Conjugation of Lipoic Acid	190
4.4.3 – Preparation of GNPs for Polyconjugation.....	193
4.4.4 – 5nm GNP16-18 Characterisation	196
4.4.4.1 – CPS Characterisation of GNP16-18.....	196
4.4.4.2 – Determining Loading of GNP16-18.....	197
4.4.4.3 – TEM Analysis of GNP16-18.....	198
4.4.5 – <i>SMN2</i> Splicing	200
4.4.5.1 – 18nm GNP Splicing	200
4.4.5.2 - <i>SMN2</i> Splicing with 20nm, 10nm and 5nm GNP	203
4.4.5.3 – Triplicate GNP16-18 <i>SMN2</i> Splicing.....	206
4.4.6 – GNP16-18 in Adenovirus Splicing	208
4.4.7 – Discussion	211
4.4.8 – Conclusion	213
4.5 – Experimental	214
4.5.1 – General Protocols	214
4.5.1.1 – Standard Ligand Exchange Protocol.....	214
4.5.1.2 – Preparation of 2'OMe RNA GNP Conjugates	214
4.5.1.3 – ³² P Labelled (Hot) Transcription Protocol.....	216
4.5.1.4 – GNP <i>in vitro</i> <i>SMN2</i> Splicing.....	217

4.5.2 – Preparation of 2'OMe RNA ONs for GNPs.....	219
4.5.2.1 – 2'OMe RNA Solid Phase Synthesis.....	219
4.5.2.2 – Preparation of RNA Lipoic Acid Derivatives	221
4.5.3 – GNP Preparation and Polyconjugation	222
4.5.3.1 – Ligand Exchange of 18nm GNP-Ci	222
4.5.3.2 – Polyconjugation of 18nm GNP-BSPP with 2'OMe RNA.....	223
4.5.3.3 – Preparation of BSPP Coated 20nm, 10nm and 5nm Particles.....	226
4.5.3.4 – Preparation of Polyconjugated 20nm GNP	227
4.5.3.5 – Preparation of Polyconjugated 10nm GNPs.....	228
4.5.3.6 – Preparation of Polyconjugated 5nm LA-RNA GNP	230
4.5.4 – GNP Splicing	231
4.5.4.1 – <i>SMN2</i> Transcription	231
4.5.4.2 – Splicing of 18nm Polyconjugated GNP 200-1nM	232
4.5.4.3 – <i>SMN2</i> Transcription	233
4.5.4.4 – 20nm, 10nm and 5nm GNP10-18 Splicing	234
5 - Main Findings from this Thesis	236
5.1 – Discussion	236
5.2 – Conclusion	239
5.3 - Future Work.....	240
Appendix.....	242
References.....	243

Dissemination of Results from this PhD Project

Papers

Perrett, A. J.; Dickinson, R. L.; Krpetić, Ž.; Brust, M.; Lewis, H.; Eperon, I. C.; Burley, G. A. *Chemical Science* **2013**, 4, 257.

Lewis, H.; Perrett, A. J.; Burley, G. a; Eperon, I. C. *Angewandte Chemie (International ed. in English)* **2012**, 51, 9800–3.

Presentations

2nd UK RNA Splicing Workshop, 4th-6th February 2011, Rydal Hall, Cumbria, UK

Department of Chemistry Postgraduate Symposium - Wednesday 6th July 2011

List of Tables

Table 1.1. Specific components of U snRNPs (adapted from Schneider et al 2010)¹

Table 2.1. Splicing data for TOES constructs tested from optimisation studies taken from Owen et al 2011²

Table 2.2. Reactive groups attached to the respective strands to enable strand ligation by bioconjugation reactions

Table 2.3. All deconstructed TOES oligonucleotides sequences for solid phase synthesis.

Table 2.4. Sequence made by RNA synthesis on an ABI 394 synthesiser

Table 2.5. Sequences made by conjugation to compound 4 from RNA1_A, RNA1_C, RNA1_E and RNA1_G.

Table 2.6. Reactants and products for all 2'OMe RNA click chemistry bioconjugation reactions undertaken

Table 2.7. Approximate percentage product conversions for click chemistry reactions

Table 2.8. Percentage isolated yields for products RNA3-RNA11 and RNA15 and their starting materials

Table 2.9. Reactants and products for all pS 2'OMe RNA click chemistry bioconjugation reactions undertaken.

Table 2.10. Melting temperatures sequences determined from CD spectra.

Table 2.11. Conditions trialed to improve product conversion during Cu-catalysed Huisgen [3+2] click chemistry

Table 2.12. RNA1_A, RNA1_C, RNA1_E, RNA2_A, RNA2_B and RNA2_C prepared using the ABI synthesizer.

Table 2.13. Products formed from click chemistry of starting materials.

Table 2.14. Products RNA1_A, RNA1_C, RNA1_E, RNA2_A, RNA2_B and RNA2_C formed by solid phase synthesis using an ABI 394 synthesiser.

Table 2.15. Characterisation summary for the RNA products RNA1_A, RNA1_C, RNA1_E, RNA2_A, RNA2_B and RNA2_C prepared by solid phase synthesis.

Table 2.16.. Products RNA1_A, RNA1_C, RNA1_E, RNA2_A, RNA2_B and RNA2_C formed by solid phase synthesis.

Table 2.17. Characterisation summary for the RNA products RNA1_A, RNA1_C, RNA1_E, RNA2_A, RNA2_B and RNA2_C prepared by solid phase synthesis.

Table 2.18. Phosphorothioate products pS RNA1_A, pS RNA1_C, pS RNA1_E, pS RNA2_A, pS RNA2_B and pS RNA2_C prepared using the ABI synthesizer.

Table 2.19. Products formed from click chemistry of starting materials.

Table 2.20. pS RNA2_A, pS RNA2_B, pS RNA2_C, pS RNA1_A, pS RNA1_C and pS RNA1_E were prepared by solid phase synthesis.

Table 2.21. RNA products prepared from the NHS reaction with compound 11.

Table 2.22. RNA14 and RNA1_G prepared by solid phase synthesis.

Table 2.23. RNA14 and RNA1_G prepared by solid phase synthesis.

Table 3.1. Exon 7 inclusion levels of SMN2 splicing in the presence of RNA3-11 and GGA-O (%).

Table 3.2. Mean exon 7 inclusion of spliced products from splicing reactions using different annealing temperatures in the presence of RNA6, RNA14, RNA15, RNA2_A and GGAO at 200nM (%).

Table 3.3. Mean exon 7 inclusion of spliced products for RNA14, RNA16, GGA and GGA-O (%).

Table 4.1. Molecular weights of RNA1_I, RNA1_J, RNA2_D and RNA2_E synthesised on an ABI 394 synthesiser.

Table 4.2. Molecular weights of RNA17, RNA18, RNA19 and RNA20 and their starting materials..

Table 4.3. Particle size and strand coating of GNP1-18.

Table 4.4. Exon 7 inclusion levels of spliced products for GNP1-9, GGA-O and RNA6 (%).

Table 4.5. Exon 7 inclusion and exclusion levels and the levels of pre-mRNA remaining after 2 hours of splicing reactions with GNP10-18 and GGA-O at concentrations of 200nM, 100nM, 50nM and 10nM (%).

Table 4.6. Mean exon 7 inclusion of spliced products for GNP16-18, GGA-O and O-TO at 200nM, 100nM and 50nM.

Table 4.7. Mean percentage of Adenovirus A2 and A3 spliced to form d/s spliced products with GNP16, GNP18 and O-TO.

Table 4.8. Shows the volume of 2M NaCl added to achieve salt concentrations for the preparation of GNP1-18.

Table 4.9. RNA ONs prepared by solid phase synthesis.

Table 4.10. Summary of RNA1_I, RNA1_J, RNA2_D and RNA2_E prepared by solid phase synthesis.

Table 4.11. Products formed from RNA1_I, RNA1_J, RNA2_D and RNA2_E reaction with compound 9.

Table 4.12. Summary of RNA17-20 prepared by NHS coupling of compound 9 to RNA1_I, RNA1_J, RNA2_D and RNA2_E.

Table 4.13. RNA17-20 polyconjugated onto GNP1-9 used in this study.

Table 4.14. Volumes of 2M NaCl in 0.5x TBE, 0.015% SDS added to achieve the final concentrations used in splicing assays.

Table 4.15. Concentration of NP conjugates GNP1-9 after removing excess RNA strands.

Table 4.16. Polyconjugated GNP10-12 used in splicing analyses.

Table 4.17. Volumes of 2M NaCl in 0.5x TBE, 0.015% SDS added to achieve the final concentrations used in splicing assays.

Table 4.18. RNA ONs RNA18 and RNA19 used in the preparation of GNP13-15 and the volume of reduced 2'OMe RNA ON added to the GNP solution.

Table 4.19. Volumes of 2M NaCl in 0.5x TBE, 0.015% SDS added to achieve the final concentrations used in splicing assays.

Table 4.20. RNA ONs RNA18 and RNA19 used in the preparation of GNP16-18 and the volume of reduced 2'OMe RNA ON added to the GNP solution.

Table 4.21. Volumes of 2M NaCl in 0.5x TBE, 0.015% SDS added to achieve the final concentrations used in splicing assays.

Table 4.22. GNP1-9 conjugates tested in a SMN2 splicing reaction.

Table 4.23. GNP10-18 conjugates used in a SMN2 splicing reaction.

List of Figures

Figure 1.1 Schematic representation of alternative splicing.

Figure 1.2. Schematic representation of the splicing mechanism and machinery.

Figure 1.3. Structure of U1 snRNP with key components.

Figure 1.4. .a) Binding of SR proteins (in green) to enhancer sequences within introns (ISE) and exons (ESE). b) Diagram showing key components of SR protein SRSF-1.

Figure 1.5. Putative cross-intron interaction between an SR protein or Prp5 (green), U2AF35 (light blue) and U1-70K (cream).

Figure 1.6. hnRNPs mode of action.

Figure 1.7. Schematic representation of the disease process of SMA.

Figure 1.8. Role of SMN in the body.

Figure 1.9. a) Design of TOES candidates on exon 7 of SMN2. b) Sequence of GGA drawn 5' to 3'.

Figure 2.1. 5' to 3' sequence of TOES oligonucleotides with LNA modifications.

Figure 2.2. Map of the positions of TOES binding investigated.

Figure 2.3 a) Deconstruction of TOES oligonucleotides to create tripartite sequences for SMN2 splicing. b) Comparison of the structure of one HEG unit and three nucleotides.

Figure 2.4. Sequences of the RNA1 series and RNA2 series incorporating linker 2. ° indicates that bases are 2'OMe RNA.

Figure 2.5. Ion exchange trace for RNA1_E purification running a gradient from 0% to 100% over 80 minutes.

Figure 2.6. RP-HPLC traces for a) RNA2_A and b) RNA1_A.

Figure 2.7. 10% denaturing polyacrylamide gel of ion exchange column fractions.

Figure 2.8. 10% polyacrylamide gel.

Figure 2.9. T_m folding and melting graphs for RNA1_A samples.

Figure 2.10. 10% denaturing polyacrylamide gel of RNA1_C and RNA1_E samples.

Figure 2.11. 20% denaturing polyacrylamide gel of RNA1_G and RNA14 samples.

Figure 2.12. Sequences RNA1_G and RNA14 incorporating linkers 8. \circ indicates that bases are 2'OMe RNA.

Figure 2.13. RNA1_A, RNA1_C, RNA1_E and RNA1_G products after NHS reaction with compound 11.

Figure 2.14. 10% polyacrylamide gels of reaction samples from bioconjugation of RNA1_D and RNA2_A.

Figure 2.15. Products of Cu-catalysed Huisgen [3+2] cycloaddition reactions characterised by 10% PAGE.

Figure 2.16. 10% denaturing polyacrylamide gel of reaction samples from bioconjugation of RNA1_F and RNA2_C.

Figure 2.17. Products of all 2'OMe RNA starting materials undergoing Cu-catalysed Huisgen [3+2] cycloaddition reactions.

Figure 2.18. 10% denaturing polyacrylamide gels of 15 minute time point samples from click chemistry reactions.

Figure 2.19. 10% denaturing polyacrylamide gels of purified RNA3-RNA11 and their starting materials.

Figure 2.20. MADLI-TOF data run in negative mode for pS DNA3 where the expected phosphothioate $[M+H] = 4965.05$.

Figure 2.21. Structures of DNA phosphothioate strands synthesised on a ABI 394 DNA/RNA synthesiser.

Figure 2.22. Products formed after RNA1_A, RNA1_C and RNA1_E were reacted with compound 11.

Figure 2.23. pS 2'OMe RNA sequences synthesised on an ABI 394 synthesiser using an "aged" PADS solution as an oxidiser.

Figure 2.24. Products of the Cu-catalysed Huisgen [3+2] cycloaddition bioconjugation reactions for pS 2'OMe RNA.

Figure 2.25. 20% denaturing polyacrylamide gels of click chemistry reaction samples. **a)** Click chemistry of pS RNA1_B with pS RNA2_A, pS RNA2_B and pS RNA2_C. **b)** Click chemistry of pS RNA1_D with pS RNA2_A, pS RNA2_B and pS RNA2_C. **c)** Click chemistry of pS RNA1_F with pS RNA2_A, pS RNA2_B and pS RNA2_C..

Figure 2.26. 20% denaturing polyacrylamide gel of pS RNA6 and pS RNA7 after purification.

Figure 2.27. RNA14 and RNA6 sequences and their GGA chemistry versions, RNA16 and RNA6A, respectively.

Figure 2.28. 20% denaturing polyacrylamide gel of RNA1_G, RNA14, RNA15 and RNA16 samples.

Figure 2.29. Secondary structure prediction for RNA1 sequence.

Figure 2.30. Strands used to investigate quadruplex secondary structures written 5' to 3' with _o and _s representing 2'OMe RNA and pS RNA, respectively.

Figure 2.31. Intermolecular quadruplex structures found in nucleic acids.

Figure 2.32. CD absorption data for RNA1, RNA18, RNA19 and pS RNA12.

Figure 2.33. NMR spectras. a) RNA1. b) RNA19. c) RNA18.

Figure 2.34. Hoogsteen base pairs in a G quartet.

Figure 2.35. Tetrad/hexad structure formed by GGA repeats intermolecularly.

Figure 2.36. Gradient used for HPLC analysis of DNA and 2'OMe RNA sequences from 5-50% 0.1M TEAA in water/0.1M TEAA 80:20 ACN:water at 1 ml/minute over 20 minutes.

Figure 2.37. Products formed from Cu-catalysed Huisgen [3+2] cycloaddition reactions between RNA1 and RNA2 series strands.

Figure 2.38. Products formed from Cu-catalysed Huisgen [3+2] cycloaddition reactions between RNA1 and RNA2 series strands.

Figure 2.39. Products of the Cu-catalysed Huisgen [3+2] cycloaddition bioconjugation reactions for pS 2'OMe RNA.

Figure 2.40. Product of the Cu-catalysed Huisgen [3+2] cycloaddition bioconjugation reaction forming RNA15.

Figure 3.1. Structure of modified RNA strands used to increase pre-mRNA binding and nuclease resistance.

Figure 3.2. Structure of the key elements of an ESSENCE oligonucleotide.

Figure 3.3. RNA3-11 tripartite sequences and GGA-O used in SMN2 splicing experiments.

Figure 3.4. Graph displaying SMN2 mean splicing ratio of exon 7 inclusion to exon 7 exclusion when using RNA3-11 tripartite sequences and GGA-O at concentrations of 100nM and 50nM after two hours.

Figure 3.5. Graph displaying SMN2 mean splicing ratio of exon 7 inclusion to exon 7 exclusion when using RNA6, RNA7, pS RNA6, pS RNA7, GGA-O and GGA-S at 100nM after two hours.

Figure 3.6. pS RNA6 and pS RNA7 tripartite sequences used in SMN2 splicing experiments.

Figure 3.7. RNA14 and RNA15 sequences used in SMN2 splicing experiments.

Figure 3.8. Graph displaying SMN2 mean splicing ratios of exon 7 inclusion to exon 7 exclusion when using RNA6, RNA14, RNA15, RNA2_A and GGA-O at 200nM after two hours.

Figure 3.9. Graph displaying SMN2 mean splicing ratios of exon 7 inclusion to exon 7 exclusion when using RNA16, RNA14, GGA and GGA-O at 200nM, 100nM and 50nM after two hours.

Figure 4.1. Looping model for tripartite sequences stimulating U1 recruitment.

Figure 4.2. Schematic depiction of polyconjugation of lipoic acid functionalised 2'OMe RNA ON 17-20 to GNP.

Figure 4.3. 2'OMe RNA ONs prepared by solid phase synthesis, written 5' to 3'.

Figure 4.4. Lipoic acid conjugated 2'OMe RNA ON 17-20 after reaction with compound 9.

Figure 4.5 Agarose (2 %) gel electrophoresis of (a) RNA-GNP (18 nm) conjugates; (b) RNA-GNP (20nm) conjugates; (c) RNA-GNP (10 nm) conjugates; (d) RNA-GNP (5 nm) conjugates used in this study

Figure 4.6. Differential centrifugal sedimentation (CPS) particle size measurements showing distributions of gold nanoparticles for GNP16 (blue), GNP17 (red) and GNP18 (black) measured in **a)** 2-50 nm and **b)** 2-10 nm size range.

Figure 4.7. TEM micrographs of **a)** PEGylated GNPs (GNP18) and RNA-functionalized GNPs **b)** GNP17 **c)** GNP16 **d)** GNP16.

Figure 4.8. Particle size distribution (PSD) charts of TEM micrograph data for GNPs **a)** GNP18 **b)** GNP17 **c)** GNP16

Figure 4.9. Fold increases in exon 7 inclusion in splicing reactions with 18nm GNP1-9.

Figure 4.10. Fold increases in exon 7 inclusion in SMN2 splicing reactions using 20nm, 10nm and 5nm GNPs (GNP10-18) and GGA-O at concentrations of 200nM, 100nM, 50nM, 10nM, 1nM and 0.4nM.

Figure 4.11. Mean increases in exon 7 inclusion (depicted as a fold increase in Exon 7 inclusion) of SMN2 pre-mRNA for 5nm GNP15-18, GGA-O and O_TO.

Figure 4.12. Adenovirus mini genes used for testing splicing enhancement by GNP16 and GNP18. A) A2 adenovirus construct. B) A3 adenovirus construct with an ESE at the 5' end of the transcript.

Figure 4.13. Mean fold increases in Adenovirus D/S splicing for GNP16, GNP18 and O-TO in A) construct A2 and B) construct A3.

Figure 4.14. Structure of 5' amino modifier with MMT protecting group used in solid phase synthesis.

List of Schemes

Scheme 1.1. Formation of $m^{2,2,7}$ G in the cytoplasm after splicing and nuclear export of m^7 G cap mature mRNA.

Scheme 1.2. The mechanism of RNA splicing showing the two trans-esterification reactions that occur in the B^* and C complexes and the consensus bases involved. Dotted lines and bases (shown in black) represent introns. Coloured blocks represent exons³.

Scheme 2.1. Proposed click chemistry mechanism for labelling RNA and DNA with the structure of TBTA ligand [Diagram adapted from Bock, V.D]⁴.

Scheme 2.2. Cu catalysed Huisgen [3+2] cycloaddition reaction for tripartite sequence without HEG.

Scheme 2.3. Post synthesis modification and bioconjugation of TOES tripartite sequences. $n = 0, 1$, and 10 for HEG and $n = 1$ and 2 for abasic linker. \circ represents 2'OMe RNA (structure shown).

Scheme 2.4. Synthetic route to make linkers for use on a DNA/RNA synthesiser. a) preparation of compound 2⁵, b) preparation of compound 8⁶.

Scheme 2.5. Synthetic route to provide reactive intermediates for Cu-catalysed Huisgen [3+2] cycloaddition. a) preparation of compound 3 and b) preparation of compound 4.

Scheme 2.6. Schematic of an NHS conjugation reaction between compound 4 and RNA 1 sequences RNA1_A, RNA1_C and RNA1_E used in bioconjugation reactions (Sections 2.5.2.11 - 2.5.2.13) \circ represents 2'OMe RNA.

Scheme 2.7. Schematic of a Cu-catalysed Huisgen [3+2] cycloaddition reaction between the RNA 1 series and the RNA 2 series to produce tripartite structures RNA3-RNA11 used in splicing reactions (Section 2.5.3.9). \circ represents 2'OMe RNA.

Scheme 2.8. Highlights the synthetic route to make the abasic phosphoramidite (8). * represent chiral centres.

Scheme 4.1. Reactions between amino modified RNA1_I, RNA2_D, RNA1_J, RNA2_E and compound 9, producing lipoic acid versions of the amino modified strands written 5' to 3'. Enhancers are red and annealers are blue.

Scheme 4.2. Polyconjugation of GNP using lipoic acid functionalised 2'OMe RNA ON 17-20. Where enhancers are in red and annealers are in blue.

Abbreviations

2'OMe – 2' methoxy RNA

ASF – Associated splicing factor

BBP – Branch point binding protein

BP- Branch Point

BPS - Branch Point Sequence

BS – Binding site

BSPP - Bis (p-sulfonatophenyl)phenyl phosphine

C terminus – Carboxy terminal region of a protein

Cap – 7- methyl guanosine

CrPi – Creatine phosphate

DCM – Dichloromethane

d/s – Downstream

DTT – Dithiothreitol

ESE – Exonic sequence enhancers

ESS – Exonic sequence silencers

ISE – Intronic sequence enhancers

ISS – Intronic sequence silencers

Kglu – Potassium Glutamate

mRNA – Messenger RNA

N terminus – Amino terminal region of a protein

NE – Nuclear Extract

NHS – N-Hydroxysuccinimide

NP40 - Nonyl phenoxypolyethoxylethanol (detergent)

ON – Oligonucleotide

PAGE – Polyacrylamide Gel Electrophoresis

PK – Proteinase K

ppt – Polypyrimidine tract

Pre-mRNA – Messenger RNA before splicing

rATP – Ribose adenosine triphosphate

RRM – RNA Recognition motif

RS – Arginine/Serine repeats

SF – Splicing Factor

SF1 – Splicing Factor 1

SF2 – Splicing factor 2

SMA – Spinal Muscular Atrophy

SMN – Survival motor neurone

snRNA – Small nuclear ribonucleic acids (U1-U6)

snRNP – Small nuclear ribonuclear proteins

SR domain –Serine/Arginine repeats

SS – Splice site

TBE – Tris-Borate-EDTA buffer

TBTA -Tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine

TCEP - Tris(2-carboxyethyl)phosphine

TOES – Targeted oligonucleotide enhancers of splicing

TOSS – Targeted oligonucleotide silencers of splicing

Tri snRNP – U4, U5 and U6 small nuclear ribonuclear proteins

U2AF – U2 auxiliary factor

ULM – U2AF ligand motif

u/s – Upstream

VEGF - Vascular endothelial growth factor

1 - Introduction

It was originally proposed that a single gene would produce a single protein, but through the human genome project it was discovered that far fewer genes were present than originally thought. The low number of human genes, ~24,500, could not account for the large number of different proteins, ~100,000, revealing that single genes were giving rise to multiple proteins⁷. It was hypothesised that in higher eukaryotes, mRNA composition must be the cause for the production of multiple proteins. Eukaryotic pre-mRNA consists of protein coding (exonic) and non-coding (intronic) regions. In order for the production of functional proteins, excision of intronic regions of pre-mRNAs and the reshuffling of exonic occurs by a process known as splicing. The average human gene contains ~7 introns which need to be spliced out prior to nuclear export⁸. The excision of introns accounts for one protein product, so for multiple proteins to be produced from one gene, exonic sequences must also be removed. Removal of intronic sequences in conjunction with exonic sequences is known as alternative splicing, but how alternative splicing is regulated is not yet fully understood.

1.1 - Alternative Splicing Mechanism

From one pre-mRNA, alternative splicing can give rise to several different mature mRNA products and therefore different protein isoforms (Figure 1.1). This process is thought to be critical for the evolution of eukaryotic organisms, with approximately 94% of human protein-encoding genes containing introns and 90% of their pre-mRNAs undergoing alternative splicing².

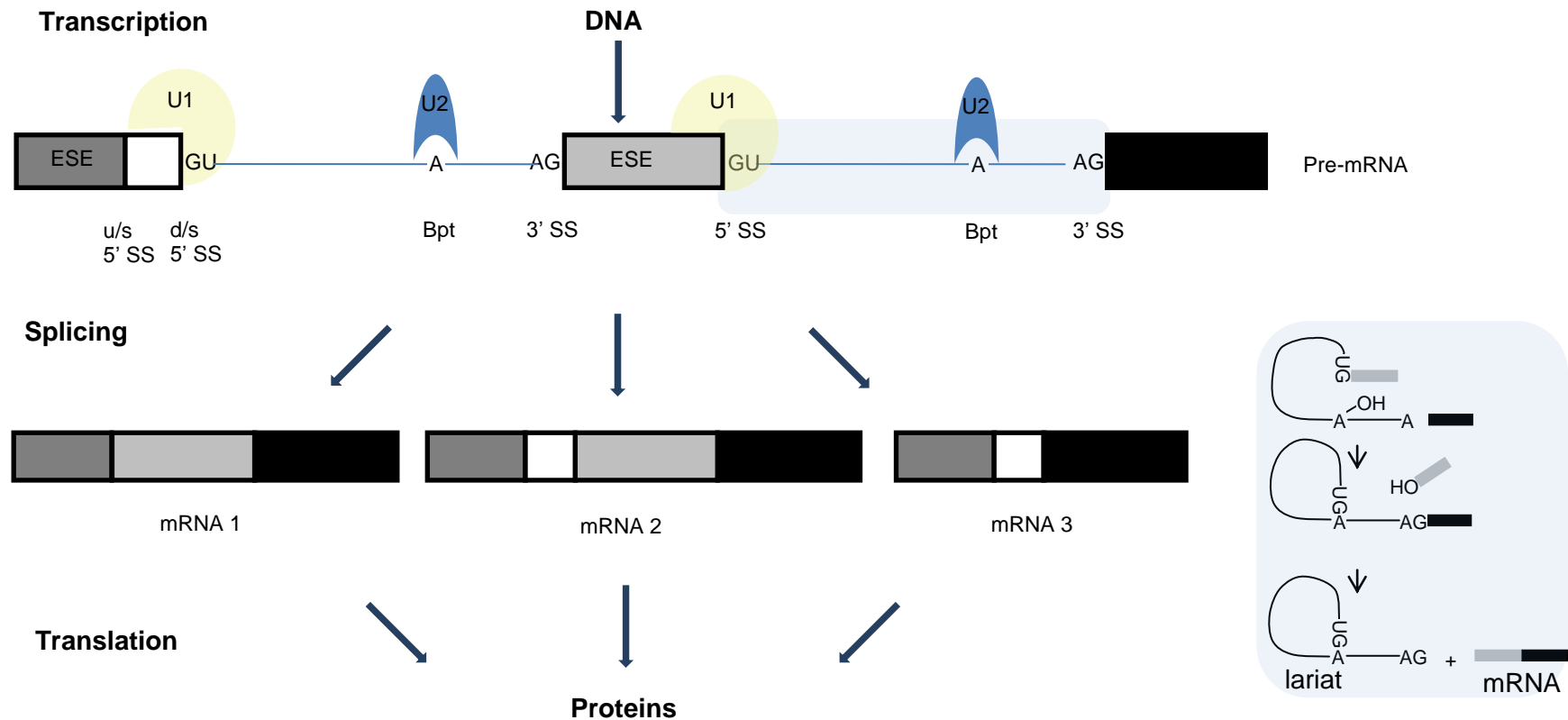


Figure 1.1 Schematic representation of alternative splicing. Rectangular blocks represent exons and lines represent introns. 5' SS and 3' SS indicate splice sites with downstream (d/s) and upstream (u/s). BP and ESE indicate branch point and exonic splicing enhancer, respectively (Section 1.1.5). The blue box shows the two transesterification reactions that form the lariat and the mRNA during the splicing reaction.

The process of intron excision and covalent assembly of exons occurs by two transesterification reactions (*Section 1.1.4*, Scheme 1.2) and is undertaken by the complex macromolecular machine known as the spliceosome. The spliceosome has uridine-rich small nuclear ribonuclear proteins (U snRNPs) named U1,U2, U4,U5 and U6. It also consists of seven Sm or Like-Sm (LSm) proteins, snRNP specific proteins, specific U-type small nuclear RNAs (snRNAs) and up to as many as 170 proteins in total (Table 1.1)^{3,9-11}. The proteins highlighted in table 1.1 are the proteins whose function in splicing in humans (higher eukaryote) has been determined. Lower eukaryotes have less introns and have more highly conserved splice site (SS) sequences, giving them a high affinity for the U snRNP, therefore not requiring enhancer and silencer proteins to determine the mRNA formed. However the proteins to catalyse splicing are very similar, with any protein with a h at the front also having a non-human protein counterpart that undertakes the same role in splicing in both higher and lower eukaryotes. The mechanism we discuss below only really considers higher eukaryotes.

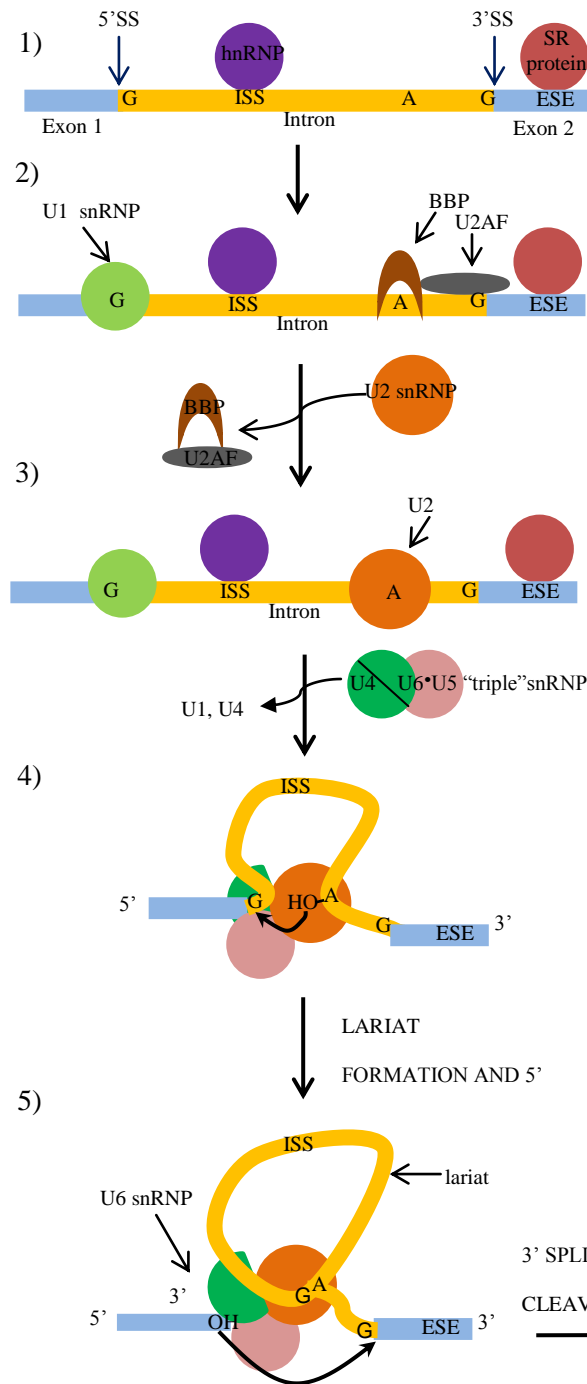
The function of the snRNPs is to regulate and catalyse the transesterification reactions. Regulation of splicing is important in determining which exons are incorporated into mature mRNA. To facilitate correct choice of splice site (SS), the spliceosome creates complexes, known as H, E, A, B, B* and C, which consist of snRNPs. Complexes H to B are required solely for splicing regulation and complexes B* and C are required for catalysis of the transesterification reactions (Figure 1.2)^{9,11}.

Table 1.1. *Specific components that have been isolated from U snRNPs in spliceosomal complexes (adapted from Schneider et al 2010)¹*

U snRNP	U1	U2	U5	U4/U6	U4/U6 U5
snRNA	U1	U2	U5	U4/U6	U4/U6 U5
Proteins specific to U snRNPs	U1-70K U1-A U1-C	U2 A' U2 B' SF3a120 SF3a66 SF3a60 SF3b155 SF3b145 SF3b130 SF3b49 SF3b14a/p14 SF3b14b SF3b10	hPrp8 hBrr2 hSnu114 40K hPrp6 hDib1 hPrp28 52K	hPrp3 hPrp4 CypH hPrp31 15.5K	hSnu66 hSad1 27K hSnu23 hPrp38

1.1.1 – H Complex Formation

The first spliceosomal complex, H complex, is a regulatory complex in which the sequences within exons and introns are recognised by enhancer or silencer proteins (*Sections 1.15 & 1.16, Figure 1.2 part 1*). Two families of proteins which recognise these sequences are serine/arginine-rich (SR) proteins and heteronuclear ribonuclear proteins (hnRNP). SR protein binding stimulates the recruitment of components for E complex formation, whereas hnRNP binding inhibits the recruitment of E complex components, therefore the balance between SR and hnRNP proteins regulates exon inclusion. The balance of these proteins is more important in weak non-consensus splice sites (SSs) as these have a low affinity for snRNP binding.



Complex H – Assembly of SR and hnRNP proteins to aid splice site choice

Complex E – Assembly of U1snRNP, Branch point Binding Protein (SF1/BBP) and U2AF35, U2AF65 and other regulator proteins.

Complex A – Binding of U2snRNP at Branch point (BP).

Complex B – The tri-snRNP U4, U5 and U6 enters and binds and U1 and U4 dissociate

Complex B* – This then allows the first transesterification to take place forming a lariat.

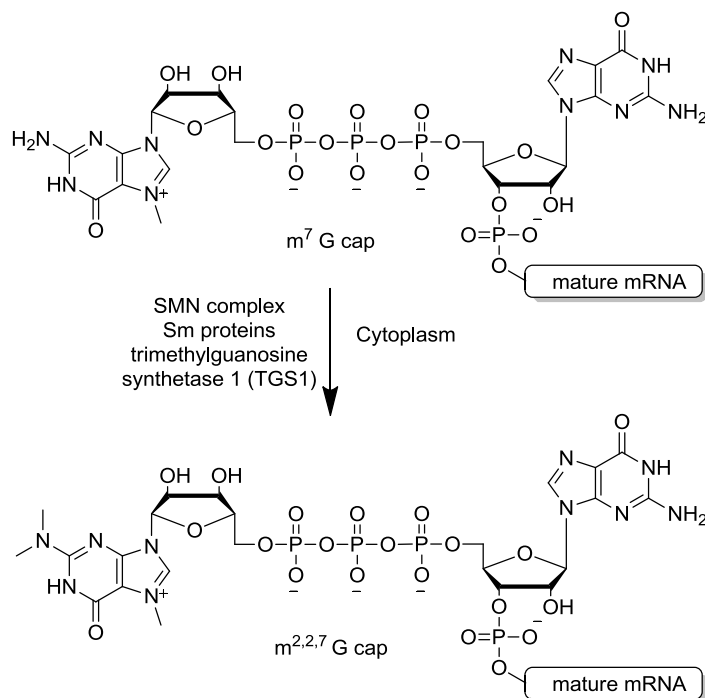
Complex C – U5 helps to keep 5' exon and 3' exon close enough for step 2 transesterification reaction to occur.

Figure 1.2. Schematic representation of the splicing mechanism and machinery.

Due to the complexity of the process only the key complexes are shown in this diagram. [Diagram adapted from Molecular Cell Biology, fifth edition]⁹.

1.1.2 – E Complex Formation

E complex formation requires the recognition of the 5' and 3' SSs and is essential for the identification of exons that will eventually be included in the mature mRNA (Figure 1.2 part 2). The exons which recruit snRNPs are determined by the corresponding SSs which are stimulated or repressed in the H complex. U1 snRNP is the first regulatory component of the spliceosome to bind the pre-mRNA, binding at the 5'SS^{12–16}. U1 snRNP consists of a U1 snRNA, a series of U1 specific proteins and seven Sm proteins^{17–21}. Sm proteins are essential for efficient methylation of the m⁷ G cap (Scheme 1.1) of a mature snRNA transcript in the cytoplasm and for nuclear transport^{22–24}. M⁷G cap is essential for stabilising mRNA and also is used as a validation cue for Sm protein binding.



Scheme 1.1. Formation of m^{2,2,7} G in the cytoplasm after splicing and nuclear export of m⁷ G cap mature mRNA.

The seven Sm proteins (B/B1, D1, D2, D3, E, F and G) each bind the Sm consensus sequence, AAUUUGUGG, present in U1, U2, U4 and U5 snRNA to form a heptameric ring around the snRNA (Figure 1.3)^{20,21,25,26}.

U1 snRNA is highly structured and this provides stem loops onto which the U1 specific proteins U1-A, U1-70K and U1-C can bind (Figure 1.3)^{10,16,19,27–29}.

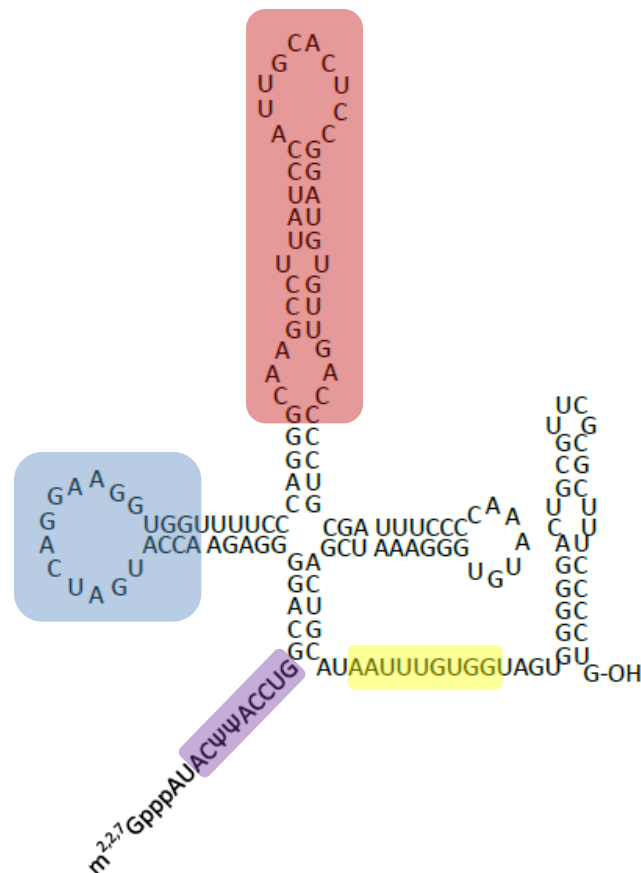


Figure 1.3. Structure of U1 snRNP with key components: stem loops I-IV, Sm binding site for Sm proteins (yellow) and 5' SS binding site (BS) (purple) which binds to the RNA. Ψ represents pseudo uridine [diagram adapted from Stark et al 2001]^{19,28}.

U1 specific protein U1-A binds stem loop II of U1 snRNA (red, Figure 1.3) through its N terminus RNA recognition motif (RRM)^{10,30}. U1-70K binds U1 snRNP with the RRM binding stem loop I of U1 snRNA (blue, Figure 1.3) and U1-70K's N terminal domain interacting with Sm proteins B/B1 and D2²⁹ which are bound to the Sm binding site (yellow, Figure 1.3). The interaction of U1-70K N terminus with Sm proteins B/B1 and D2 is essential for U1-C binding as a binding groove is created between the N terminal domain of U1-70K and the C terminus of SmD3¹⁰. This binding groove is essential for stable U1 snRNP binding to the pre-mRNA at 5'SS to a consensus sequence CAGGUAAGU (purple, Figure 1.3), enabling complex progression^{10,16}.

3'SS recognition requires the branch point sequence (BPS), a polypyrimidine tract (ppt) and the 3' SS AG. The branch point consensus sequence is UACUAAC which is poorly conserved in mammals, therefore requiring protein interaction to determine the correct branch point sequence^{31,32}. The splicing factor SF1 (previously known as BBP) is an essential protein for BPS binding, however in mammals BPS binding requires an interaction between SF1 and U2AF65^{31,33,34}. U2AF65 stabilise SF1 binding to the BPS by an interaction between the C terminal domain of the third RRM of U2AF65 and the N terminus of SF1³⁴⁻³⁹. On forming this protein-protein interaction the pre-mRNA bends enabling the U2AF ligand motif (ULM) of U2AF65 to come into contact with the 3'SS and to enable the RS domain of U2AF65 to come into contact with the BPS³⁹⁻⁴¹ further stabilising the complex formation.

The small component of the U2AF heterodimer, U2AF35, is an essential component for stimulating E complex formation as SR proteins recruit U2AF35 to the 3' SS (*Section 1.1.5.2*)^{42–44}. U2AF35 then binds the ULM of the U2AF65 which stabilises the pre-mRNA binding of the heterodimer U2AF65/U2AF35 and allows progression into A complex^{39,42–46}.

1.1.2 – A Complex Formation

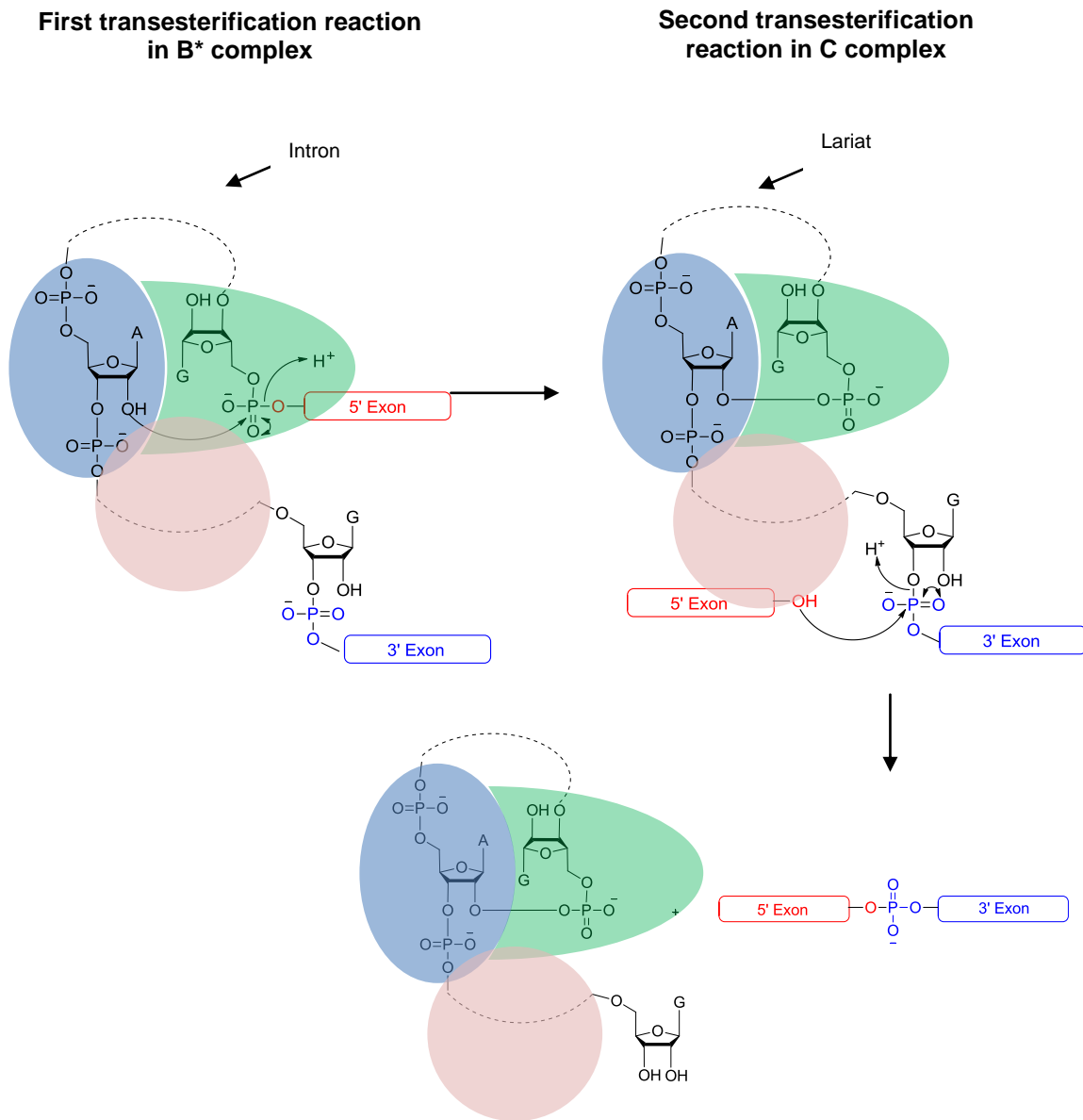
A complex formation is the first ATP-dependent step, allowing the previous complexes to be checked prior to further spliceosome assembly (Figure 1.2 part 3). The splicing complex is validated by removing SF1 binding to the BPS and replacing this with U2 snRNP binding^{33,47,48}. The removal of SF1 requires ATP as it requires disrupting the protein-protein interaction between SF1 and U2AF65 for a protein-protein interaction between U2AF65 and U2 snRNP specific splicing factor (SF) SF3b155⁴⁹. To further aid the removal of SF1, SF1 binding to BPS is proposed to be disrupted by a rearrangement at the BPS caused by UAP56, which is an U2AF65 associated RNA dependent helicase^{50,51}. U2 snRNP binding act as a regulatory step in the spliceosome formation, ensuring a correct BPS is identified in E complex formation. Essential U2 snRNP specific regulatory proteins SF3a/b, bind to sequences upstream (u/s) of the BPS to stabilise U2 snRNP binding (Table 1.1)^{52–56}. The BPS-U2 snRNP binding is further stabilised by the RS domain of U2AF65 and through interactions with SF3b155⁴¹. A cross-intron interaction between Prp5 and both U1 and U2 snRNP completes A complex formation which can then progress into B complex (Figure 1.2 part 4)⁵⁷.

1.1.3 – Splicing Complexes B, B* and C

Formation of the B complex requires the recruitment of the U4/U6.U5 tri-snRNP along with tri-snRNP specific proteins (Figure 1.2 part 4)^{3,58}. The B complex is the last point at which the spliceosome can detect whether correct SSs have been chosen to determine whether or not spliceosome progression will continue. This validation is important to ensure that cryptic SS aren't chosen over natural SS, producing functional proteins, minimising disease pathways. This regulation consists of base pair interactions between the U4/U6 snRNP and the U2 snRNP along with U5 snRNP replacing U2AF binding at the 3' SS and interacting with the 5' SS^{3,58,59}. B complex regulation is further aided by the binding of U4 snRNP to U6 snRNP which masks the catalytic important region of U6 snRNP, inhibiting U6 snRNP binding to the 5' SS. This inactivation of U6 snRNP by U4 snRNP during recruitment is to ensure that the correct SS have been chosen before the pre-mRNA splicing reaction. In order for the spliceosome to become active, a conformational rearrangement is required in which U1 and U4 snRNPs are removed from the spliceosome^{3,60}. U4 snRNP is removed by the DExD/H box protein Brr2, in association with Snu114, which destabilise the interaction between U4 snRNP and U6 snRNP, releasing the catalytic important regions of U6 snRNP⁶⁰⁻⁶². U6 snRNP and Prp28 (DEAD box protein) then destabilise U1 snRNP^{60,63}. The destabilised U1 snRNP is then displaced by U6 snRNA with the consensus sequence of U6 snRNA ACAGAG binding the 5' SS forming the B* complex^{3,60,63}.

The splicing B* complex (Figure 1.2 part 4) provides an appropriate molecular environment for the nucleophilic 2'OH of the adenosine nucleotide branch point (BP) to attack the phosphate phosphorus of the 5' exon between the first intron base and the first 5' exon base (Scheme 1.2)³. This first transesterification reaction results in a free 5' exon and a lariat-3' exon construct (Scheme 1.2)³. For the second transesterification reaction to occur, C complex requires the recruitment of further splicing factors to associate with it³. C complex then brings the 5' exon close to the 3' exon allowing the second transesterification reaction to occur in step 2, resulting in the release of the lariat (Figure 1.2 part 5 and Scheme 1.2)³. The U2, U5 and U6 snRNPs are then freed, allowing them to partake in further splicing reactions³.

The process of alternative splicing is highly regulated and only when the spliceosome reaches the B* complex does the actual excision of sequences occur. In pre-mRNA constructs with less well defined SSs (non-consensus SSs) there are other factors that are required to regulate SS selection in addition to snRNPs. These other factors include SS strength, enhancer proteins, silencer proteins and other sequences within pre-mRNA.



Scheme 1.2. The mechanism of RNA splicing showing the two trans-esterification reactions that occur in the B* and C complexes and the consensus bases involved. Black dotted lines and bases represent introns. Coloured blocks represent exons³. U2, U6 and U5 are represented by the blue, green and pink shaped respectively.

1.1.4 – Enhancers

Enhancer proteins bind RNA sequences and increase incorporation of particular exons into mature mRNA by enhancing the recognition of proximal SS signals (Figure 1.4 a). In pre-mRNA there are binding sites for enhancers in both introns and exons called intronic splicing enhancers (ISE) and exonic splicing enhancers (ESE) respectively (Figure 1.4) ⁶⁴. The best defined set of enhancers are serine/arginine-rich (SR) proteins which possess arginine-serine-rich repeats in the arginine-serine (RS) domain at the C terminal and at least one RNA Recognition Motif (RRM) at the N terminal ^{65–69}. The RS domain binds to other proteins and the RRM binds RNA (Figure 1.4b) ^{65–69}. SR proteins can enhance splice sites that are hundreds of nucleotides away giving rise to difficulties in predicting SS selection ⁷⁰.

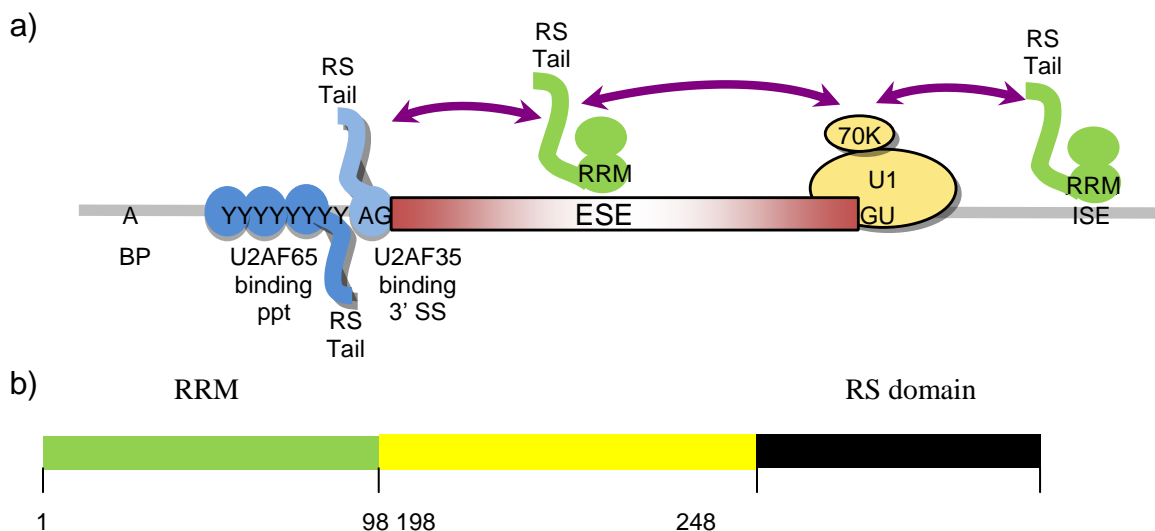


Figure 1.4. a) Binding of SR proteins (in green) to enhancer sequences within introns (ISE) and exons (ESE). b) Diagram showing key components of SR protein SRSF-1. RRM is an RNA recognition motif and is essential for RNA binding. RS domain is a series of arginine and serine repeats which form protein-protein interactions facilitating splicing.

1.1.4.1 – 5'SS Enhancement by SR Proteins

Strong SSs are sites which are highly complementary to the ten nucleotide consensus SS binding site in U1 snRNA, with weaker SSs being less complementary. If completely complementary, 25 hydrogen bonds will be formed between a SS and the U1 snRNA binding site, but a weak SS will only hydrogen bond to 4 or 5 of the binding site bases, forming a maximum of 15 hydrogen bonds if 5 C-G bonds are formed between the SS and the U1 snRNA binding site⁷¹. U1 snRNA will therefore have a higher affinity for stronger 5' SSs, which will stabilise the interaction between them, allowing spliceosome assembly to progress^{12,72–76}.

Most eukaryotic 5' SSs are weak, resulting in weak interactions for these sites with U1 snRNA. It is therefore surmised that U1 snRNP binding alone does not determine SS selection⁷⁵. Binding of SR proteins to ISEs and ESEs within the pre-mRNA compensates for the low affinity of U1 snRNP for the 5' SS. SR proteins SRSF-1 and SRSF-2 (previously SF2/ASF and SC35) recruit U1 snRNP to the 5'SS by interacting with the U1 snRNP specific protein U1-70K (Figure 1.4)^{77–79}. SR protein SRSF-2 has also been implicated in a cross-intron interaction between U1-70K and U2AF35 indicating that enhancing at the 5'SS also helps stabilise spliceosome formation at the 3'SS (Figure 1.5)^{78,80,81}.

In systems where there are two alternative 5'SSs in an exon, most SR proteins shift splicing to the downstream (d/s) SS both *in vitro* and *in vivo*^{77,82–86}. In the case of SRSF-1, U1 snRNP is recruited to both 5' SSs equally and the shifting effect occurs because the SS in closest proximity to the 3' SS is usually selected (Section 1.1.7)^{77,83}.

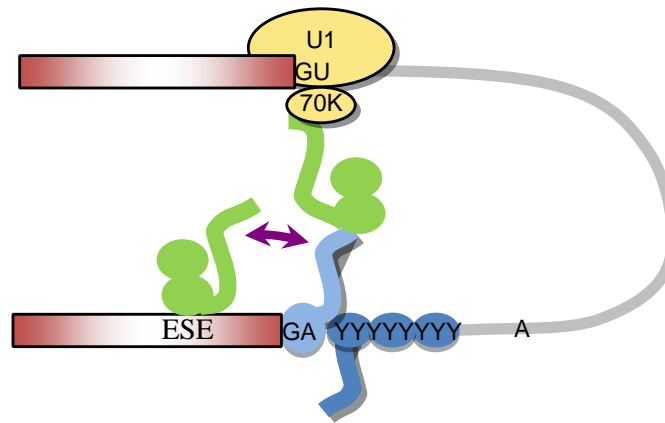


Figure 1.5. Putative cross-intron interaction between an SR protein or Prp5 (green), U2AF35 (light blue) and U1-70K (cream). U2AF35 is bound to the 3' SS and U2AF65 (dark blue) is bound to the ppt. Rectangular blocks represent exons and grey line indicates introns.

SR proteins are not only thought to be important for early spliceosome assembly. They have also been implicated with U4/U6 U5 tri-snRNP recruitment⁸⁷. U6 and U5 snRNPs are further stimulated by SR proteins, after recruitment, with the RS domain of the SR protein stabilising the binding of U6 and U5 to the 5'SS and sequences u/s of the 5'SS, respectively^{88,89}.

1.1.4.2 – 3'SS Enhancement by SR Proteins

As with the 5' SS, SR proteins are involved in recruiting spliceosome components to the 3' SS. SR proteins interact with the RS domain of U2AF35, recruiting U2AF65 to a weak 3' SS^{78,90–97}. U2AF65 binding is then able to recruit U2 snRNP to the weak 3' SS allowing splicing to occur (Figure 1.4)⁴⁹. U2AF65 also has an RS domain which may form an interaction with ESE's in the absence of U2AF35⁹⁰. SR proteins have also been implicated in U2 snRNP recruitment during A complex

formation where RS domains interact with the BPS. These interactions can either accelerate or stabilise U2 snRNP binding to the BPS by neutralising the electrostatic repulsion between U2 snRNP and the pre-mRNA BPS⁹⁸. As with SR proteins at the 5' SS, SR proteins at the 3' SS are not only associated with early stage spliceosome assembly, they have also been shown to be capable of recruiting both U2 and U6 snRNP to some pre-mRNA in the absence of U1 snRNP, indicating that high concentrations of SR proteins can compensate for loss of U1 snRNP^{98–100}. In a system with two alternative 3' SSs, SR proteins cause splicing to occur at the 3' SS closest to the 5'SS (*Section 1.1.7*)^{101,102}.

SR proteins can block binding of splicing inhibitors (silencers) to pre-mRNA by recognition of binding sites^{103,104}. An example of this is c-src exon N1 in which the binding site for both SRSF-1 and hnRNP A1 are within the same 11 nucleotides (*Section 1.1.6*)¹⁰⁴.

SR proteins have been found to act as inhibitors when bound to an intronic splicing silencer sequence, inhibiting the recognition of a 3' SS^{105,106}. However, moving this sequence from an intron to an exon creates an ESE sequence, indicating that SR proteins' effect on spliceosome recruitment is dependent on binding sequence position within the pre-mRNA^{105,106}.

1.1.5 – Silencers

Silencers are proteins which bind to pre-mRNA sequences, preventing recognition of splice sites by the spliceosome and stopping incorporation of exons into mature mRNA. Like enhancers, silencers are found in both introns and exons. Trans-

acting factors that bind exonic splicing silencers (ESSs) and intronic splicing silencers (ISSs) are a family called heteronuclear ribonuclear proteins (hnRNP) ⁶⁴.

1.1.5.1 – SS Inhibition by hnRNP

hnRNPs, when bound to pre-mRNA, reduce U1 snRNPs binding affinity for a 5'SS, resulting in excision of exonic RNA ¹⁰⁷. Binding of hnRNP A1 and hnRNP H to a pre-mRNA in a 5'SS competition assay causes destabilisation of U1 snRNP from the d/s 5' SS, shifting splicing to the u/s 5' SS ^{108–111}. It has also been suggested that hnRNP H can bind to 3'SSs, to inhibit U2AF35 binding ¹¹². hnRNP L regulates splicing differently by blocking the transition between an exon defined complex to an intron defined complex rather than exerting an effect on a SS (*Section 1.1.9*) ^{113–116}. hnRNPs can also act together to modulate 5' SS selection by looping out or smothering the pre-mRNA and inhibiting the recognition of the 5' SS by the spliceosome (Figure 1.6) ¹¹⁷. Even hnRNPs, which are known silencer proteins, can, in certain situations, aid SR protein recognition of SSs by looping out sections of introns ^{117,118}.

1.1.6 – Proximity Effect

If two 5'SSs are more than 40 nucleotides (nts) apart and of similar SS strength then splicing occurs at the 5'SS closest to the 3'SS ^{77,119–121}. This suggests there is a correlation between the distance between SSs and SS selection. If the two 5'SSs are less than 40nts apart and of similar SS strength, splicing occurs at both 5' SS ¹¹⁹. This indicates that RNA diffusion may be responsible for SS selection when both 5'SSs are bound by U1 snRNP, with the 5'SS closest to the 3' SS more likely to encounter the 3'SS and therefore be spliced ¹²⁰.

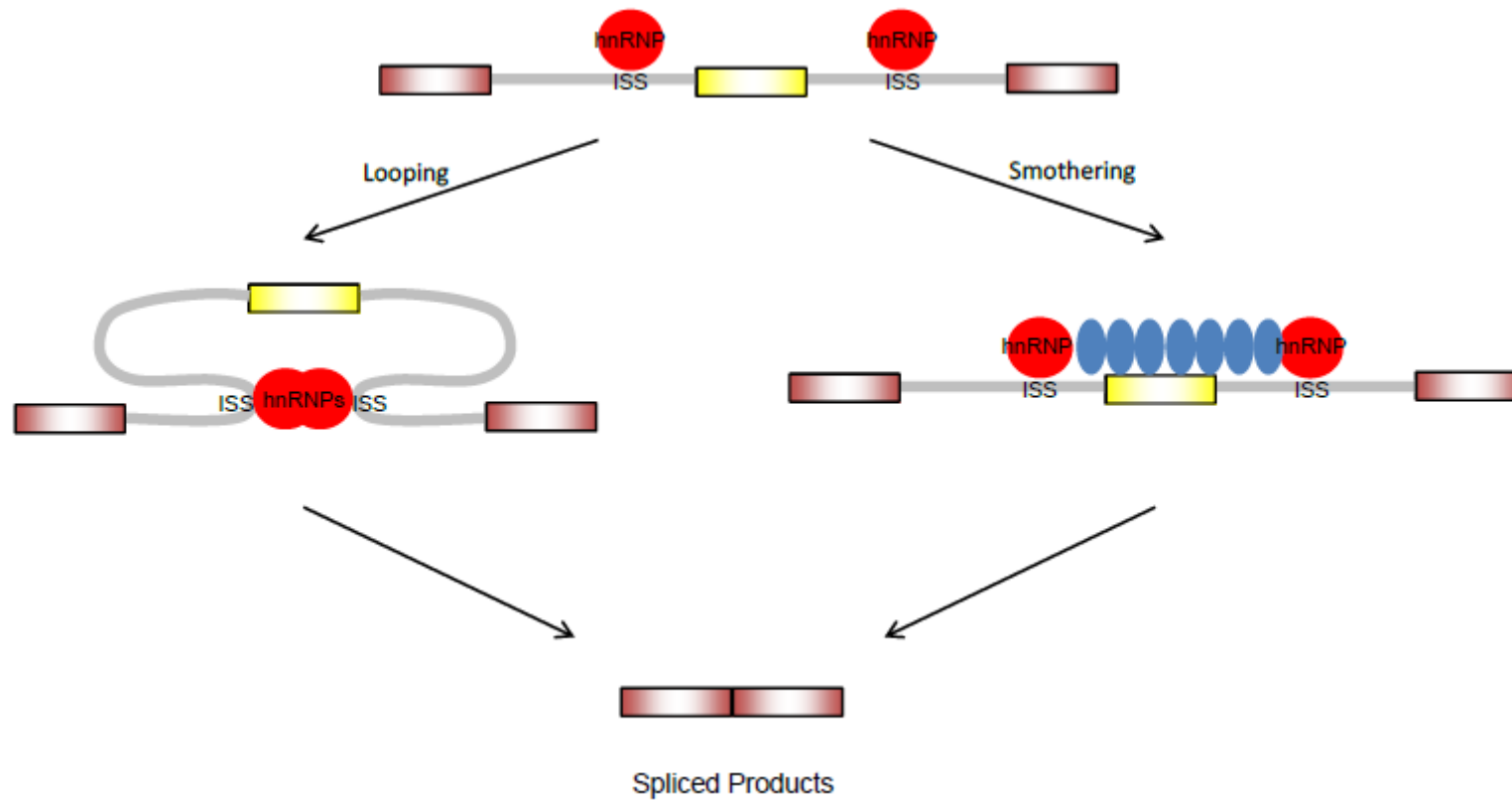


Figure 1.6. *hnRNPs mode of action: hnRNPs bind to the ISS regulatory sites on the pre-mRNA and cause looping out or smothering of exonic pre-mRNA, to inhibit splice site recognition by the spliceosome machinery. Rectangular blocks represent exons and grey lines represent introns. Other unidentified silencer proteins are shown in blue.*

1.1.7 – 3'SS Selection

5' SS selection is purely due to the binding of U1 snRNP, whereas 3'SS selection is more complex. The 3' SS consists of a BPS, a ppt and a AG dinucleotide at the 3'SS. ppt strength is key to whether a pre-mRNA requires proteins to promote U2AF65 binding¹²². In pre-mRNA with a weak ppt, strong binding is required between SF1 and BPS and between U2AF35 and the 3'SS to stabilise U2AF65 binding (*Section 1.1.2*)^{39–44,123}. Strong ppt however, can be bound by U2AF65 without stabilisation from SF1 or U2AF35¹²². In the absence of U2AF35, binding of U2AF65 to enhancer-dependent pre-mRNA does occur, however no splicing of the pre-mRNA is observed, therefore recruitment of U2AF65 is not presumed to be the sole function of U2AF35^{39,124}.

1.1.8 – Exon Definition to Intron Definition

Exon definition is the connection between the 5' SS and 3' SS of an exon during SS recognition. Exon defined complexes occur on pre-mRNA with introns longer than ~200 nts and exons shorter than ~300 nts (which is the case for most mammalian pre-mRNA), enabling exons to be determined in the midst of thousands of intron nts^{1,125}. When SR proteins bind ESE sequences it causes a network of protein-protein interactions, bridging both the U1 snRNP at the 5' SS and U2 snRNP at the BPS, forming an exon defined complex (Figure 1.4 a)^{1,125,126}. Exon defined complexes are thought to be the reason why so few cryptic SSs are used during splicing, as the use of cryptic SS would lead to non-functional proteins. As splicing occurs across introns, exon defined complexes need to be converted into cross-intron defined complexes (Figure 1.5)^{1,125}. The transition between these

two complexes determines which SSs are used and which exons are included in the final pre-mRNA¹. The point at which cross-exon defined complexes switch to cross-intron defined is unknown, however inhibitory factors PTB, RBM5 and hnRNP L have been shown to block the transition from exon defined to intron defined complex after U1 and U2 snRNP binding^{1,113–115,125}. This would suggest the transition occurs during A complex formation, however exon defined complexes have been shown to associate with the U4/U6 U5 tri-snRNP in the B complex¹. Depending on when the exon definition to intron definition transition occurs, inhibitors could stop the transition by blocking the recruitment of the U4/U6 U5 tri-snRNP¹. It is therefore not known when this transition occurs and even whether cross-exon defined complexes have to be broken for cross-intron defined complexes to occur. However, what is known is that whenever the switch between exon definition and intron definition occurs it is responsible for regulating whether an exon is included or skipped from the final pre-mRNA¹.

In order for the spliceosome to function efficiently, a delicate balance of splicing factors and cues is required. The main pre-mRNA sequences required for SS recognition are the 5'SS, 3'SS, BPS and the ppt. These sequences alone are not always strong enough to stimulate splicing at weaker splice sites, as the consensus SS are not always present, but enhancer sequences can aid SS recognition. SR proteins (enhancer proteins) can bind to enhancer sequences to enable recognition of weak splice sites, but hnRNP proteins (silencer proteins) can bind silencer sequences to prevent splice site recognition. Whether a SS is used or not depends on whether hnRNP or SR proteins exert the greatest effect. With a

system so complex, the disruption of one or more of the components involved could potentially shift SS selection, leading to aberrant splicing events and resulting in the production of non-functional proteins and therefore disease pathways¹²⁷.

1.2 – Alternative Splicing and Disease

Aberrant splicing events are linked to numerous diseases since even the smallest splicing error that adds or removes a single nucleotide, disrupts the open reading frame (ORF) of a gene. This results in an aberrant mature mRNA product and a non-functional protein which can ultimately cause the onset of disease^{128,129}. In an average human gene there are 7 introns with ~3,300 nucleotides in each intron and 8 exons with ~150 nucleotides per exon, which correlates to approximately 29,000 nucleotides being transcribed per gene into pre-mRNA¹²⁸. A large proportion of alternative splicing events undergo cell-specific regulation, hence different cell types can have different susceptibilities to specific splicing mutations¹²⁸. The majority of mutations that disrupt splicing are caused by single nucleotide polymorphisms (SNPs) which result in either complete exon skipping, use of pseudo splice sites, or intron retention in the mRNA product¹²⁸. In the case of pseudo splice sites and intron retention, the introduction of premature stop codons in the mRNA causes the translation of truncated protein isomers, which typically results in nonsense mediated decay and loss of function¹²⁸.

There are two types of mutation; cis-acting and trans-acting. Cis-acting mutations affect the pre-mRNA, but not the splicing machinery. When a cis-acting mutation

disrupts the binding sequence of an enhancer or silencer, it can, through aberrant splicing events, lead to the formation of a mutant mRNA isoform which can in turn lead to disease¹²⁸. An example of a cis-acting mutation is atypical cystic fibrosis which is caused by aberrant pre-mRNA splicing which produces a non-functional cystic fibrosis transmembrane conductance regulator (CFTR) protein¹²⁸. A trans-acting mutation is one in which the splicing machinery itself is affected, which has the potential to affect splicing of all pre-mRNA and prevent formation of important proteins. Such a mutation is responsible for Spinal Muscular Atrophy which is a disease caused by inadequate quantities of the essential protein, SMN, in spinal cord motor neurons.

1.3 – Spinal Muscular Atrophy (SMA)

SMA is the most common genetic cause of childhood mortality. 1 in 40 people are carriers and 1 in 6,000 births are affected worldwide from SMA^{130,131}. It is an autosomal recessive disorder, which in 95% of cases *SMN1* is absent and in 5% of cases it is mutated⁸. SMA, caused by a lack of functional SMN protein, results in the loss of spinal cord motor neurons resulting in skeletal muscle denervation and subsequent paralysis of voluntary muscles^{8,128,132}. The genes responsible for motor neuron survival are *SMN1* and *SMN2* which give rise to SMN protein^{8,128,130–133}. These two genes are nearly identical apart from 11 nucleotides, however one nucleotide in exon 7 that causes the different spliced products observed between *SMN1* and *SMN2*¹³³. *SMN2* spliced products do not usually include exon 7, resulting in a protein with severely decreased stability (Figure 1.7)^{8,128}. The cause of this frequent exon 7 skipping during splicing is a thymine at position +6 in exon 7

(a cytosine is at this position in *SMN1*)^{8,128,131–135}. This thymine not only disrupts an ESE sequence for SRSF-1 but also creates an ESS binding site for hnRNP A1, with both contributing to the loss of exon 7 inclusion in most *SMN2* mRNA^{133,135,136}. On the occasions that exon 7 of *SMN2* pre-mRNA is not excised, stable SMN protein is produced, since the thymine does not affect translation.

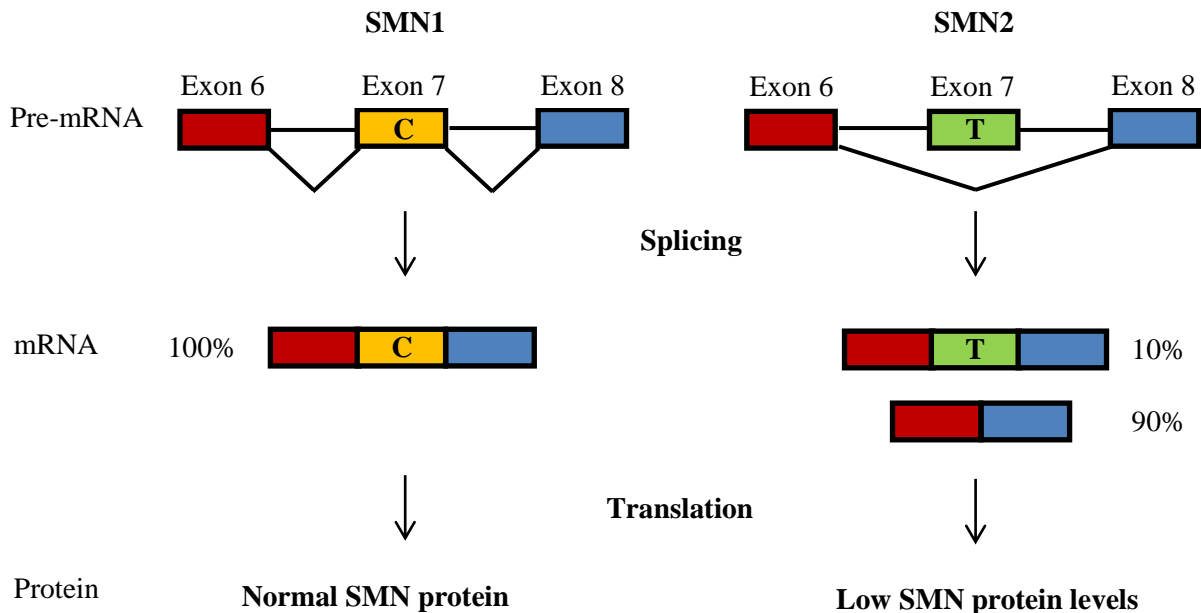


Figure 1.7. Schematic representation of the disease process of SMA. The survival motor neuron genes *SMN1* and *SMN2* differ in exon 7 by eleven nucleotide. The majority of *SMN2* splicing products lack exon 7 due to the T base at position +6 of exon 7. Therefore, loss of *SMN1* results in reduced SMN levels in most tissues.

SMA has three different degrees of severity in children with type I being the most severe and type III patients having milder symptoms¹³⁷. Children affected by type I SMA, which accounts for 50%-60% of SMA sufferers, usually die within 1-2 years, whereas children with type III can live a relatively normal life^{137,138}. The severity of

SMA depends on the number of *SMN2* copies present^{128,137}. This is because the higher the copy number of *SMN2* a patient has, the greater the amount of functional SMN protein will be produced, which can help to compensate for the lack of *SMN1*. Therefore, the lower the *SMN2* copy number, the more severe the symptoms. Most type I SMA patients have only one or two copies of *SMN2*, most type 2 patients have three copies of the gene and the majority of type three patients have three or four copies¹³⁷.

SMN1 exons 1-7 encode a 294-amino acid protein that joins with at least seven other proteins, Gemins 2-8, to form the SMN complex which is required for Sm protein assembly on the snRNA (U1, U2, U4, U5) in the cytoplasm^{8,128,130,132}. Sm proteins are a heptameric ring of proteins which stabilise snRNA and recruit other factors to enable nuclear transport forming snRNPs^{8,22-24}. Once in the nucleus, the snRNP is complexed with more proteins, preparing it for pre-mRNA splicing (Figure 1.8)^{8,22-24}. SMN not only facilitates Sm protein assembly on snRNA but also stringently controls it to inhibit Sm protein assembly on non-snRNA¹³⁰.

The discovery that SMA severity inversely correlates to the number of *SMN2* genes means that stimulating exon 7 inclusion in *SMN*, therefore producing more protein, could prove an attractive therapeutic strategy for the treatment of SMA.

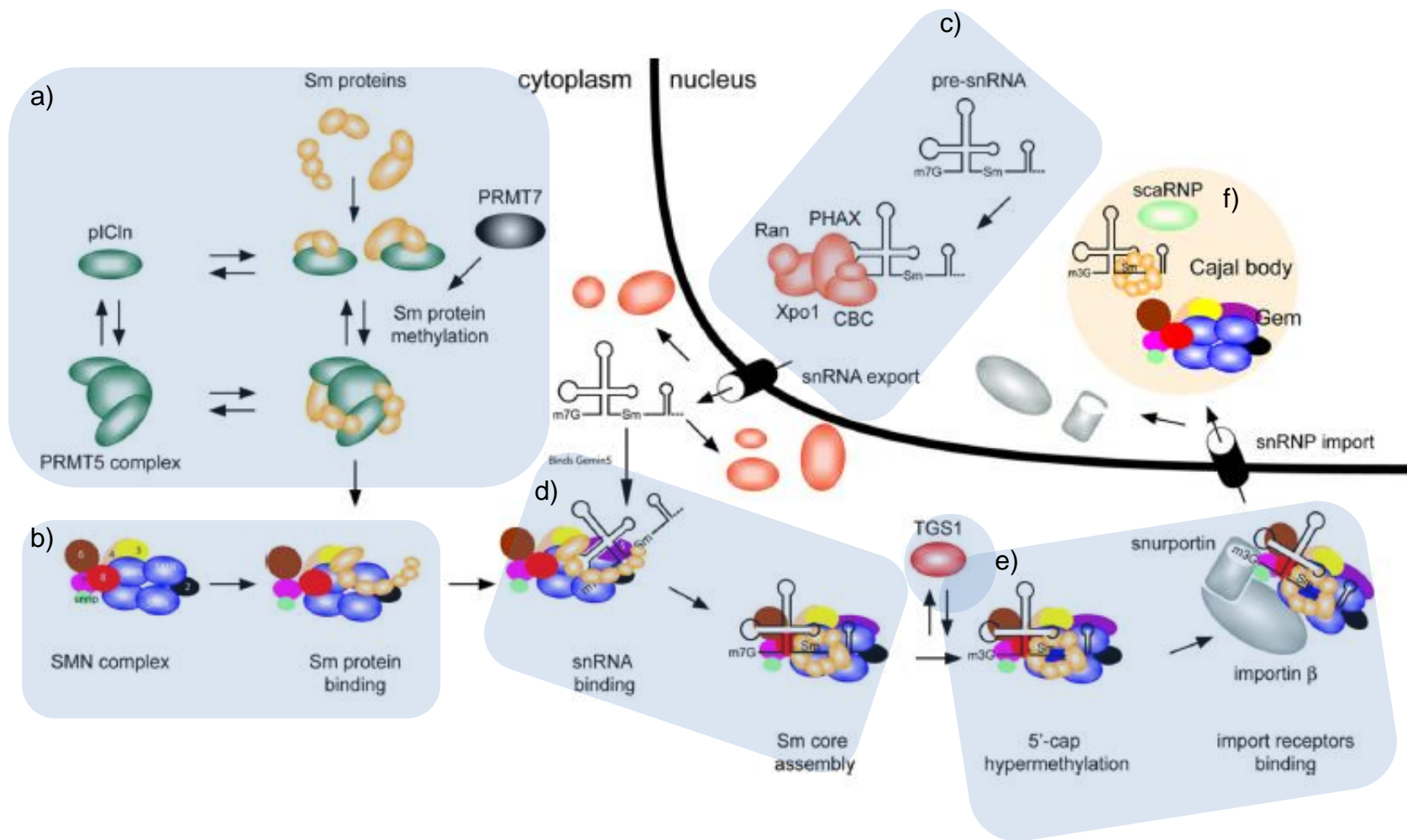


Figure 1.8. *snRNPs are composed of small nuclear RNA (snRNA), seven Sm proteins and several specific proteins. Survival motor neuron (SMN) assembles Sm proteins onto the snRNAs to produce active snRNP. a) In the cytoplasm, three Sm proteins are methylated before binding to the SMN complex. b) The SMN complex is shown as an oligomer as it has been suggested that oligomerisation is crucial for SMN function. c) snRNA is transcribed in the nucleus and then exported into the cytoplasm by proteins e.g. Ras-related nuclear protein (RAN), exportin (XPO1)cap-binding complex (CBC) and phosphorylated adaptor for RNA export (PHAX).d) The SMN complex puts the Sm proteins onto the snRNA. e) The 7-methylguanosine (m7G) cap of the snRNA is hypermethylated by trimethylguanosinesynthetase 1 (TGS1), allowing the binding of importin and snurportin, which mediates the transport of the SMN complex and the snRNP into the nucleus. f) In the nucleus, the snRNPs undergo further maturation with the SMN complex in Cajal bodies [Diagram taken from Burghes et al 2009⁸].*

The problem with trying to stimulate exon inclusion in any disease system is that ESE, ESS, ISE and ISS sequences are hard to predict using current bioinformatics techniques. Binding of RNA strands to ESE and ISE sequences will be detrimental to exon inclusion, with the opposite being true for ESS and ISS sequences. Therefore in order to deduce an optimal location for stimulating exon inclusion using an RNA binding approach, large quantities of RNA oligonucleotides would need to be produced. However, probably only one or two of the many hundreds of strands synthesised would actually stimulate exon inclusion. Due to the costs involved with synthesising RNA oligonucleotides, this approach is undesirable, especially in disease pathways where the cause of the switch in splicing patterns, which leads to the disease, is not so well understood. There are many methods which have been shown to stimulate exon 7 inclusion in *SMN2*, such as using anti-sense oligonucleotides (ASO) and PNA-peptide derivatives (ESSENCE) (*Section 3.2*), however the strategy investigated as part of this PhD project is termed Targeted Oligonucleotide Enhancers of Splicing (TOES).

1.4 – TOES

TOES were originally designed as a splicing tool to modify splicing outcomes, in a variety of disease systems, where mutation within exons caused suboptimal levels of proteins¹³⁹. The design principle of TOES was to include an annealing region, complementary to the mutated exon, and an enhancer region able to mimic an ESE and encourage protein binding (Figure 1.9 a).

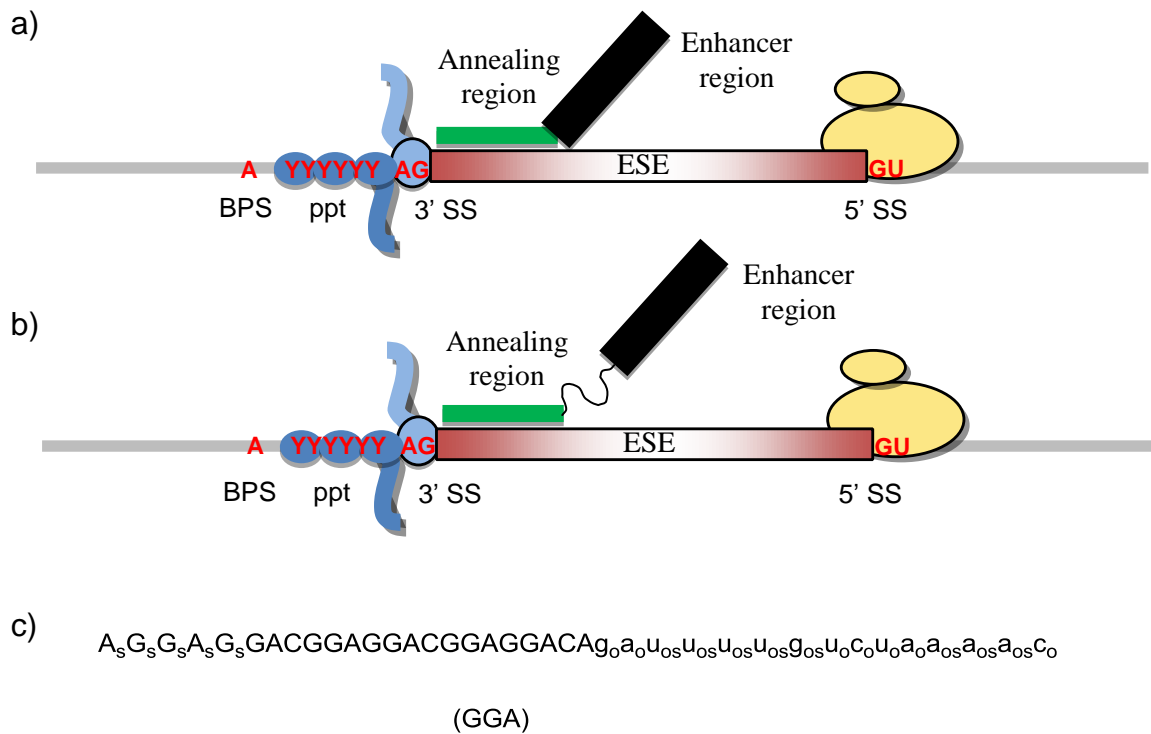


Figure 1.9. a) Design of TOES candidates on exon 7 of SMN2 and b) shows the design of the tripartite sequence. Annealing region is green, enhancer region is black and black curved line is a bioconjugation group and linker. Grey lines represent introns and the red box exons. U2AF65 is dark blue, U2AF35 is light blue and U1 snRNP is yellow. b) Sequence of GGA drawn 5' to 3'. Capital letters are the enhancer and lower case letters the annealer. s and o represent phosphorothioate and 2'OMe RNA, respectively¹³⁹.

The known ESE sequence used in TOES is a GGA repeat, GGAGGAC, which is known to bind to SR protein SRSF-1 and has been shown to stimulate 5' SS recognition (Section 1.1.5.1)¹⁴⁰. SMN2 is an ideal disease system to investigate the TOES strategy as stimulation of exon 7 inclusion is critical for SMN protein levels (Figure 1.9).

GGA oligonucleotide (Figure 1.9 b), when tested in both *in vitro* and *in vivo* resulted in an increase in relative proportion of exon 7 inclusion to levels of 1.5 - 2.3 at 50nM and 250nM respectively compared to exon 7 inclusion levels of 0.25 for the control, which contained no GGA. An annealer only version (NT) of GGA was used as a control to determine whether the increase seen in exon 7 inclusion, occurred due to the presence of the enhancer or just the binding of the annealer over the ESS¹³⁹. The NT version only achieved inclusion levels of 0.25 – 0.7 for 50nM and 250nM respectively proving that the enhancer is required for efficient exon 7 mRNA inclusion.

TOES were originally designed to be used to rectify splicing defects in disease systems in which splicing is defective on a specific exon. This however would require screening of TOES constructed wholly by solid phase synthesis which would be an expensive process (Figure 1.9 a)). This problem can be overcome by synthesising enhancer and annealer sequences separately, then bioconjugating them together to form a bifunctional oligonucleotide (tripartite sequence, Figure 1.9b)). The benefits of making tripartite sequences, rather than TOES, are that shorter (and therefore less expensive) strands can be synthesised, but more importantly once a particular annealer or enhancer sequence is synthesised it can be used to make a multitude of different tripartite sequences. Therefore, if a bioconjugation group is tolerated between an annealer and enhancer sequence, it could provide a cost-effective, powerful method for screening for bioactive TOES.

In a TOES system, an enhancer sequence binds SRSF-1, which then recruits U1 snRNP to the 5'SS¹³⁹. In this type of monovalent system only one TOES

construct can bind to one pre-mRNA strand at any given time. Increasing the number of GGA repeats in an enhancer sequence enhances splicing in this type of system, since the more GGA repeats there are; the more opportunities there are for SR protein binding. Extension of the enhancer region to include more GGA repeats has been carried out in the TOES system, with results showing that exon 7 inclusion levels increased as the number of GGA repeats were increased². The disadvantage of using longer oligonucleotides is that they are more expensive to produce and more difficult to put into cells due to their higher molecular weight and increased negative charge. By using a multivalent system, large numbers of enhancer and annealer strands can be effectively brought together by conjugating them to a surface. This will enable key components, such as pre-mRNA and SR proteins, to be brought together more effectively because there will be a larger number of binding sites for them in one particular place. Gold nanoparticles (GNPs) will be investigated for use as a multivalent system surface for binding enhancer and annealer sequences. This has not been investigated before for splicing, despite the potential for such a multivalent system being more efficacious than a monovalent system.

1.5 – Aims

Aim 1: To assess whether flexible non-RNA linkers can enhance the splicing activity of TOES. Longer oligonucleotides have been shown to be beneficial for correcting splicing defects (Section 2.1) however increasing the length using RNA increases number of binding sites on TOES for proteins therefore it is unknown whether the length or the extra protein binding is responsible for the increased activity. By using non-RNA linkers to increase the length we can

determine whether length or extra protein binding sites are responsible for the increased TOES activity. A further reason for investigating non-RNA elements within TOES is that when investigating other disease systems there will be a requirement to adapt the TOES sequence. It would therefore be beneficial to incorporate a conjugation group into the non-RNA part allowing annealer and enhancer sequences to be bioconjugated together from lab stocks to effectively screen annealers and enhancers pair activities in different disease pathways.

Aim 2: To determine whether a multi-valent display of TOES components enhance splicing. By increasing the number of protein binding sites at a particular place, using Au nanoparticles, we can determine what effect increasing the number of protein binding sites has on TOES activity.

1.6 – Hypothesis

That non-RNA elements can enhance splicing activity either in a tripartite construct or a multi-valent display on a nanostructure.

This thesis will be split into three major sections:

1. Synthesis and bioconjugation of TOES tripartite sequences (*Chapter 2*).
2. *SMN2* splicing assay with the TOES tripartite sequences. Investigating whether the bioconjugation group is tolerated in the TOES tripartite sequences (*Chapter 3*).
3. Preparation and *SMN2* splicing assay for multivalent nanoparticle systems. Determining whether Au nanoparticle (GNP) size and coating will effect exon 7 inclusion levels (*Chapter 4*).

Chapter 2 – Preparation of tripartite oligonucleotides

2 – Introduction

Since the original paper on TOES in 2003, further work has been undertaken in pursuit of finding the optimal TOES constructs ¹³⁹. As the TOES contain both an annealer and an enhancer sequence, there are many different factors that have to be considered when attempting to optimise the system. To optimise the annealer sequence, the strength of binding and the position of the binding required for optimal exon 7 inclusion have to be considered. To optimise the enhancer sequence, the number of protein binding sequences, the sequence of the enhancer and the chemical structure of the bases have to be considered.

2.1 – Aims

TOES are essentially made up of an annealer and an enhancer strand. By separating the two and inserting a bioconjugation group and linkers, increasing the diversity of the annealer and enhancer pairs, the aim is to attempt to correct the splicing defects in well studied disease systems and to use this approach to investigate other disease systems where suboptimal levels of a desired protein are produced. By using this approach the aim is not only to save on costs, but also on time, by bioconjugating strands already in stock. Synthesis of these strands will require optimisation as the longer TOES tripartite sequences will have up to 10 units of HEG making the strand very flexible, potentially affecting solid phase synthesis efficiency. To produce these TOES tripartite sequences, the two strands require conjugating together. This will be carried out using Cu-mediated click chemistry. At the time the work was started, no click chemistry reactions had been carried out on RNA or 2'OMe RNA, therefore reaction and purification conditions were unknown.

2.2 – TOES Optimisations

2.2.1 – TOES Annealer Optimisations

TOES 1 (also known as GGA) was the first example of a TOES which contained six repeats of a known enhancer sequence (GGA) and a +2 to +16 binding site for exon 7^{139,140}. When annealing an oligonucleotide to an exon, especially one in which there is a binding site for a silencer protein (hnRNP A1) as in exon 7 *SMN2* (Section 1.3), the strength of binding to that site could play a pivotal role^{135,136}. To investigate binding strength, LNA oligonucleotides were prepared because LNA derivatives have been shown to increase affinity for RNA over 2'OMe RNA (Figure 2.1)². LNA versions, TOES 9, 10, 13 and 17 (Figure 2.1) were less active than the 2'OMe/pS counterparts, TOES 1 and 3, indicating that increasing binding to +2 to +16 in exon 7 of *SMN2* decreases exon 7 inclusion levels especially in type II SMA patient fibroblast cells (Table 2.1). RNA oligonucleotides require protection from degradation and so pS, 2'OMe RNA and LNA derivatives have been used as caps at the 5' and 3' end of oligonucleotides².

The decreased inclusion levels seen with increased binding strength of TOES is potentially due to the TOES only being required for early stage spliceosome assembly. The binding of TOES to the +2 to +16 regions in exon 7 may be too close to the 3'SS and require displacement for further spliceosome assembly. Moving the annealing position of the TOES away from the 3' SS in exon 7 could potentially increase the exon 7 inclusion levels that the LNA TOES stimulate in *SMN2*. To investigate this, several different annealing positions between +2 and +46 of exon 7 were tested (Figure 2.2).

A_sG_sG_sA_sG_sGACGGAGGACGGAGGACAg_oa_ou_{os}u_{os}u_{os}u_{os}g_{os}u_oc_ou_oa_{os}a_{os}a_{os}C_o
(TOES 1)

A_sG_sG_sA_sC_sGGAGGACAg_oa_ou_{os}u_{os}u_{os}u_{os}g_{os}u_oc_ou_oa_{os}a_{os}a_{os}C_o
(TOES 6)

A*G*G*A*G*GACGGAGGACGGAGGACAg_oa_ou*_uu*_ug*_uc_ou_oa_oa*a*a*c_o
(TOES 9)

A*G*G*A*C*GGAGGACGGAGGACAU_ou_ou*_uu*_ug*a*_uu_ou_ou_og*u*c*_u
(TOES 12)

C*A*A*G*A*AGAAGAAGCUg_oa_ou*_uu*_ug*_uc_ou_oa_oa*a*a*c_o
(TOES 14)

c*c*_uu_oc_oc_ou_ou_oc*_uu*_uu*_ug_oACAGGAGGCAGGAGG*C*A*G*G*A
(TOES 18)

A*G*G*A*C*GGAGGACGGAGGACAc_oa_oc*c*_uu*_uc*_uc_ou_ou_oc_ou*_uu*_u
(TOES 21)

A_oG_oG_oA_oG_oG_oA_oC_oG_oG_oA_oG_oG_oA_oC_oG_oG_oA_oG_oG_oA_oC_oA_og_oa_ou_ou_ou_og_ou_oc_ou_oa_oa_oa_oC_o
(TOES 27)

A_sG_sG_sA_sC_sGGAGGACGGAGGACAg_oa_ou_{os}u_{os}u_{os}u_{os}g_{os}u_oc_ou_oa_{os}a_{os}a_{os}C_o
(TOES 3)

A_sA_sC_sC_sA_sGACGACAGACGAAAg_oa_ou_{os}u_{os}u_{os}u_{os}g_{os}u_oc_ou_oa_{os}a_{os}a_{os}C_o
(TOES 7)

A*G*G*A*C*GGAGGACGGAGGACAg_oa_ou*_uu*_ug*_uc_ou_oa_oa*a*a*c_o
(TOES 10)

AGGACGGAGGACGGAGGACAg_oa_ou*_uu*_ug*_uc_ou_oa_oa*a*a*c_o
(TOES 13)

A*G*G*A*C*CGCGGACCGCGGACAg_oa_ou*_uu*_ug*_uc_ou_oa_oa*a*a*c_o
(TOES 17)

A*G*G*A*C*GGAGGACGGAGGACAU_ou_oa*a*g*g*a*_uu_og_ou_og*a*g*c_o
(TOES 19)

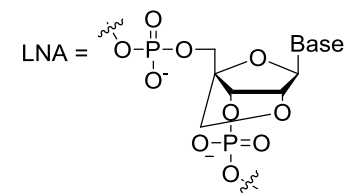


Figure 2.1. 5' to 3' sequence of TOES with LNA modifications. Capital letters indicate the enhancer and lower case letters indicate the annealer. The _s, * and _o indicate the phosphothioate backbone, LNA base and 2'OMe RNA, respectively.

Table 2.1. Splicing data for TOES constructs tested from optimisation studies taken from Owen et al 2011²

Strand	<i>In vitro</i> exon 7 inclusion level	<i>In vivo</i> exon 7 inclusion level
TOES 1	1.9	7.58
TOES 3	1.78	4.93
TOES 6	1.3	2.95
TOES 7	1.21	2.66
TOES 9	1.58	3.47
TOES 10	1.55	2.85
TOES 12	1.16	1.46
TOES 13	1.18	1.39
TOES 14	1.11	1.38
TOES 17	1.15	0.96
TOES 18	0.56	0.06
TOES 19	0.12	0.02
TOES 21	0.91	0
TOES 27	1.41	5.84

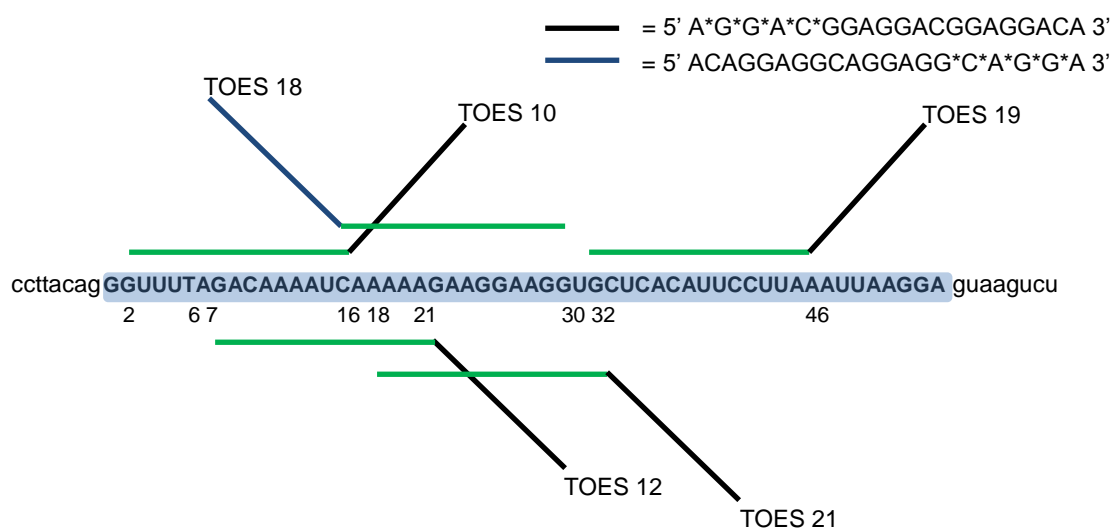


Figure 2.2. Map of the positions of TOES binding investigated. Blue box indicates exon 7 of SMN2, * indicates LNA base, lower case letters indicate introns, green lines indicate annealers complementary to the pre-mRNA and black and blue lines indicate enhancers (sequences shown) [diagram adapted from Owen et al 2011]².

In these tests the levels of exon 7 inclusion decreased as the distance of the annealing site from the 3'SS increased (Table 2.1). The data indicates that +2 to +16 binding in exon 7 is optimal for maximum enhancer contribution to exon 7 inclusion in *SMN2*. Switching the enhancer to the 3' end of the TOES, as in TOES 18, causes the enhancer to exert very little effect. The decreased effects seen for TOES 18 and 21 oligonucleotides may be due to the sequences binding across a known Tra2 β binding site between positions +22 to +28¹³¹. TOES 12 binding could also be interrupting the Tra2 β binding site as TOES 12 binds +7 to +21 therefore could sterically disfavour Tra2 β binding at sites +22 to +28. TOES 19 may be binding so close to the 5' SS that it is affecting spliceosomal factors binding, or it could be too far away from the 3' SS to enable enhancement.

None of the attempts to optimise the annealing portion of the TOES for *SMN2* splicing proved successful; therefore enhancer optimisations were investigated.

2.2.2 – TOES Enhancer Optimisations

The sequence GGAGGA is known to bind enhancer protein SRSF-1¹⁴⁰. To investigate the number of GGA repeats required for efficient exon enhancement, six, five and three repeats of GGA were tested (Figure 2.1).

The ability of the TOES to enhance exon 7 inclusion decreased when the number of GGA repeats were reduced from six (TOES 1) to five (TOES 3) to three (TOES 6) (Table 2.1). This suggests either multiple repeats are required to recruit SRSF-1 or that the repeats allow multiple copies of SRSF-1 to bind². The sequences CAGACG, GAA and CCGCGGA are also known to bind SRSF-

1^{2,140}. Therefore enhancers were constructed with CAGACG, GAA and CCGCGGA repeats to investigate what effect they would have on splicing in *SMN2* and how this compares to the effect seen with the GGAGGA sequence (Figure 2.1)².

The new enhancer sequences used in TOES 7, 14 and 17, gave rise to exon 7 inclusion levels that were lower than those seen with the GGAGGA sequences in TOES 1, 3 and 10 (Table 2.1). Therefore, the GGAGGA enhancer sequence is most preferable for stimulating exon 7 inclusion in *SMN2*.

Of all the TOES constructs synthesised up to this point, TOES 1 had given rise to the highest exon 7 inclusion levels. In an attempt to further improve TOES 1, the RNA enhancer sequence was changed to 2'OMe RNA to create TOES 27 (Figure 2.1).

Changing the enhancer sequence to 2'OMe RNA led to a decrease in exon 7 inclusion levels, however the levels were still high in comparison to some of the other optimisation oligonucleotides tested (Table 2.1).

The optimisation studies undertaken have revealed that for *SMN2*, +2 to +16 is the optimal binding position for TOES and weak binding (predicted to be ~31°C) gives higher exon 7 inclusion levels. The investigations also showed that six repeats of the GGA motif in the enhancer sequence provided the greatest exon 7 inclusion levels for *SMN2*. Also enhancer sequence activity has been shown to be dependent on the strand's structure, with RNA pS giving rise to higher inclusion levels than 2'OMe RNA. After all the modifications tested, TOES 1 (the original construct) seemed to provide the greatest enhancement of exon 7 inclusion in an *SMN2* system. However, not all systems will behave the same

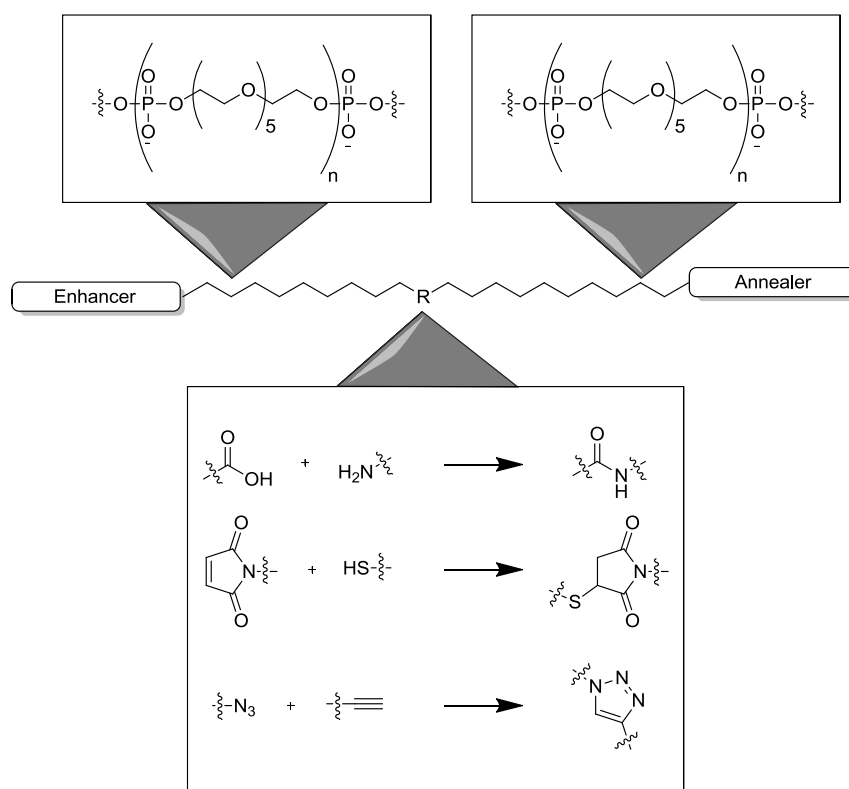
as *SMN2* or have ESEs and ESSs in the same places, therefore different splicing defects will require oligonucleotide screening and optimisation to determine the best annealer/enhancer pair for a particular disease pathway. To efficiently screen for enhancer/analyzer pairs a fast effective route into these compounds needs to be addressed. To investigate routes into these compounds different bioconjugation groups will be considered.

2.2.3 – Tripartite Sequence

Deconstructing the TOES constructs into the two main parts, the enhancer and the analyzer, requires the insertion of a bioconjugation group to facilitate the reconstruction of the TOES tripartite sequences. The insertion of a bioconjugation group between two strands for use in splicing has not been investigated, and hence the biological activity of this bioconjugation group on splicing is unknown. To determine whether a bioconjugation group's proximity to the enhancer and analyzer strands affects enhancement of exon 7 inclusion, hexaethylene glycol (HEG) linkers were used to space the bioconjugation groups away from the strands. One HEG unit is equivalent in length to three nucleotides, but unlike nucleotides it has no charge and has a lower molecular weight (Figure 2.3 b). HEG linkers were used because they create flexibility within strands which may increase the ability of the enhancer sequences to interact with the 5'SS of exon 7 to enable splicing (Figure 2.3 a). HEG was also chosen because it had previously been shown to be inert in splicing reactions when incorporated into an intron¹⁴¹. Other reasons for using HEG are that it is neutral, non-toxic and hydrophilic, which is ideal for use *in vivo* (Figure 2.3). Proteins will not bind to HEG linkers and will therefore not interfere with enhancer binding nor decrease

strand flexibility (which makes using HEG preferential to using oligonucleotides because they would bind proteins). A range of HEG repeats were incorporated into different tripartite sequences to determine optimal spacing. With the spacer confirmed, investigations into the appropriate bioconjugation group were undertaken.

a)



b)

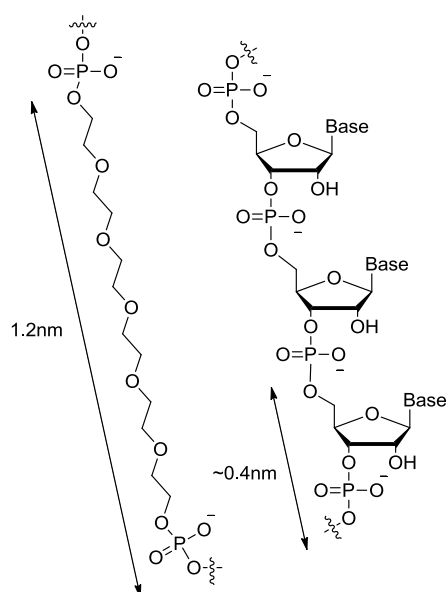
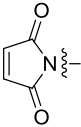


Figure 2.3. a) Deconstruction of TOES to create tripartite sequences for SMN2 splicing. b) Comparison of the structure of one HEG unit and three nucleotides.

2.2.4 – Bioconjugation Chemistry for Tripartite Sequences

To improve the TOES use as a general splicing tool, the choice of bioconjugation group is essential. The bioconjugation group not only has to link the two halves together but has to be tolerated in between the two strands for the system to maintain activity. To reduce the effect of the presence of the bioconjugation group, small groups were chosen (Figure 2.3, Table 2.2).

Table 2.2. Reactive groups attached to the respective strands to enable strand ligation by bioconjugation reactions

5' AGGAGGACGGAGGACGGAGGACA–R ¹ 3' + 5' R ² –GAUUUUGUCUAAAAC 3'		
(Enhancer)		(Annealer)
R ¹	R ²	Conjugation Reaction
NH ₂	CO ₂ H	Amide coupling
	SH	Thiol-maleimide coupling
N ₃	≡	Click chemistry ¹⁴²

The reason these conjugation groups were chosen as potential reactive groups for the bioconjugation reaction is because they all have been shown to conjugate two nucleic acid strands together in the literature. This shows that the reactive groups are not only compatible with biomolecules, but that the relevant reactive groups can be attached to the nucleic acid strands.

2.2.4.1 – Amide Coupling

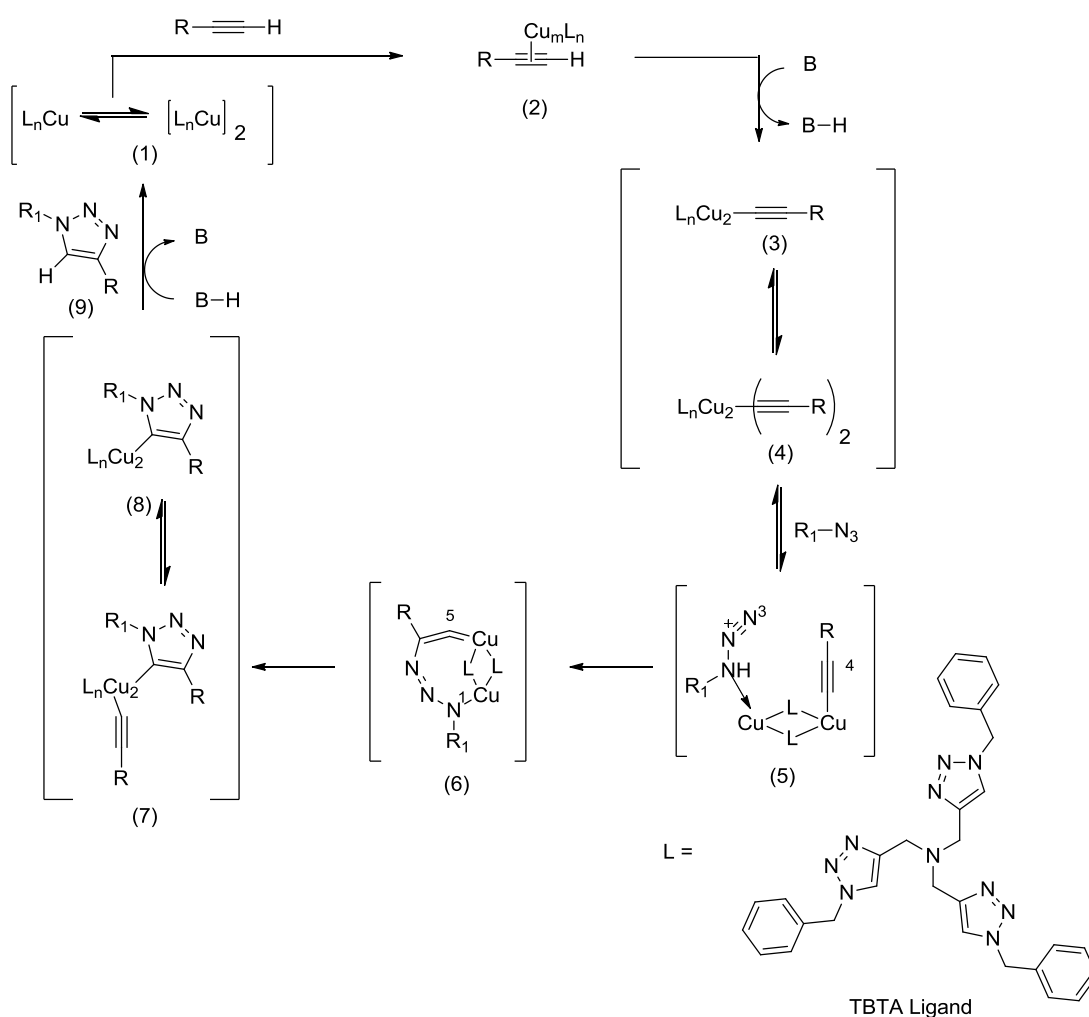
Amide coupling is the only reaction in which the reactive groups can be inserted during solid phase synthesis without post-synthesis modification. The carboxylic phosphoramidite can be attached to the 5' end of a strand during solid phase synthesis. The amino modifiers come as either phosphoramidites or on resin allowing both the reactive groups to be conjugated to the individual strands, enabling the bioconjugation reaction. Amide coupling reactions on a DNA-carboxylic acid derivative with a DNA-amine using templated synthesis have been shown to achieve yields of around 80% with the reaction without a template achieving ~46% after 12 hours at RT¹⁴³. However the storage of the activated carboxylic acid could potentially be a problem and therefore other bioconjugation methods were investigated.

2.2.4.2 – Thiol-Maleimide Coupling

To insert the reactive groups for a thiol-maleimide coupling, the 5' annealer end can be modified with the thiol phosphoramidite during the solid phase synthesis reaction. The enhancer side would require deprotection and a post-synthesis modification to insert a maleimide group onto the 3' end. This reaction has been carried out before on DNA strands in the presence of a template sequence giving a yield of between 60-90%¹⁴⁴. The reaction however does not occur in the absence of a template. This limits the range of reactions that this conjugation reaction can be used for especially as some of the template strands would need to be approximately 80nts in length for bioconjugations to occur¹⁴⁴.

2.2.4.3 – Copper Catalysed Huisgen [3+2] Cycloaddition (Click Chemistry).

Cu catalysed 1,4 cycloadditions have been the reactions of choice for labelling DNA and RNA due to their fast reaction times and quantitative conversions¹⁴⁵. The proposed click chemistry mechanism involves coordination between Cu and the alkyne followed by the cycloaddition (Scheme 2.1).



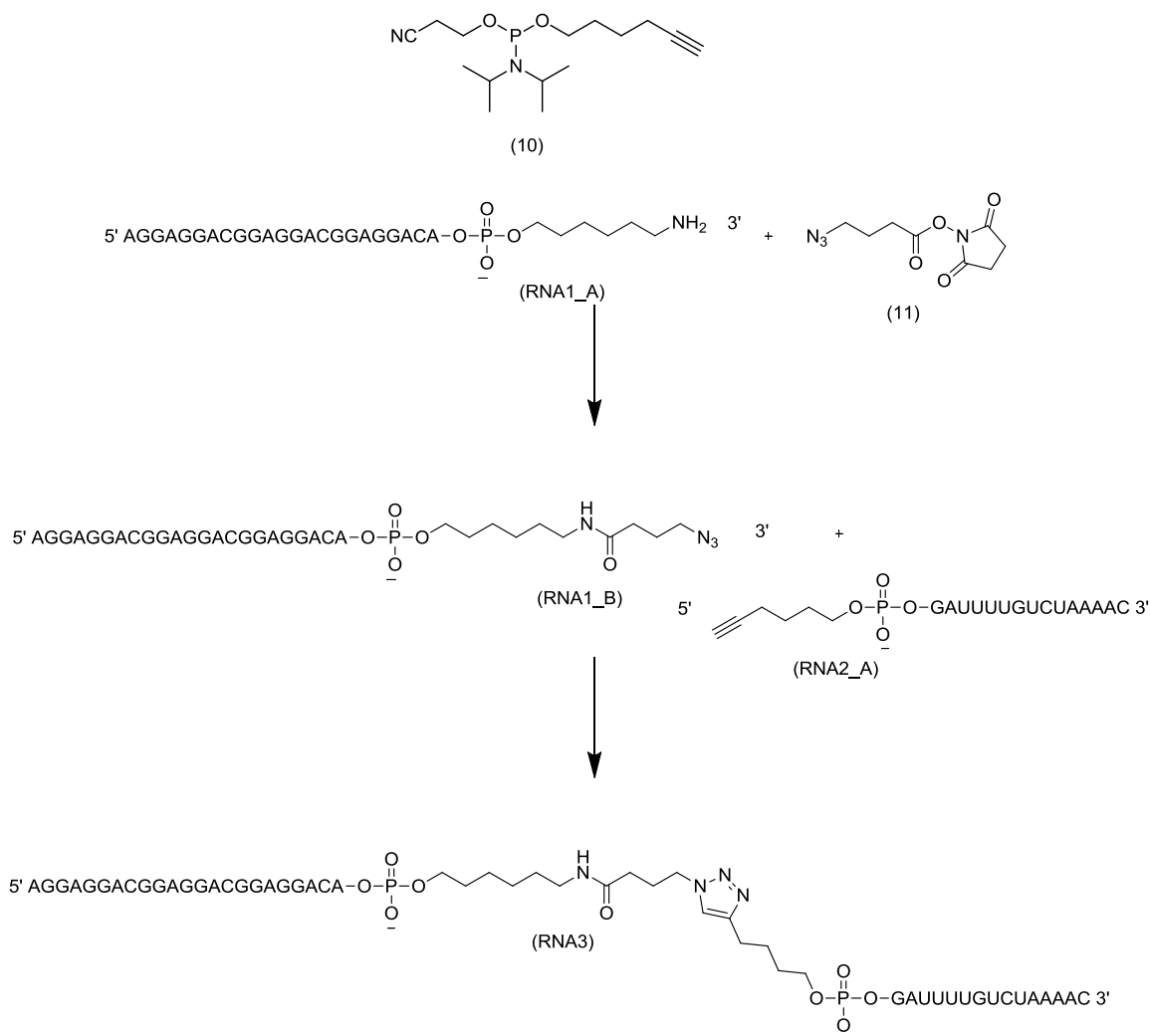
Scheme 2.1. Proposed click chemistry mechanism for labelling RNA and DNA with the structure of TBTA ligand [Diagram adapted from Bock, V.D]⁴.

The major concern with using the Cu-catalysed reaction with biomolecules is that Cu produces OH radicals which can degrade DNA and RNA strands^{146,147}. However this problem is negated by the use of an appropriate ligand (TBTA, Scheme 2.1, Compound 1), not only enabling increased reaction rates, but also

inhibiting OH radical production^{146,147}. Compound **1** then forms a Cu acetylide species through the π bond forming compound **2** (Scheme 2.1). Compound **2** could undergo a click chemistry reaction with an azide however the activation barrier for an azide reacting with compound **2** is 11.7kcal/mol higher than the reaction of an azide with compounds **3** or **4** and therefore disfavoured (Scheme 2.1). In aqueous solutions the rearrangement from compound **2** to compound **3** is exothermic and occurs rapidly, with the coordinating Cu weakening the alkyne C-H bond allowing it to be broken without the addition of a base. The actual structure of compound **4** is not known however a second Cu species shown in compound **4** is required for the activation of the azide forming a Cu acetylide-azide complex (Compound **5**, Scheme 2.1). This activation of the azide allows the nucleophilic attack of the Cu-azide N3 into C4 of the Cu-alkyne forming a metallocycle (Compound **6**, Scheme 2.1). The metallocycle allows N1 and C5 to be close enough in proximity for the formation of compound **7** to occur via the addition of the N1 lone pair to the π^* orbital of the C5 (Scheme 2.1). Addition of a proton and dissociation from the Cu complex produces compound **9** (Scheme 2.1).

To insert the reactive intermediates for click chemistry the 5' end of one of the strands can be modified with an alkyne phosphoramidite during solid phase synthesis (Scheme 2.2, compound **10**). The 3' strand can be modified with an amino CPG which on deprotection can undergo a N-Hydroxysuccinimide (NHS) reaction with a NHS azide (Scheme 2.2, compound **11**)¹⁴⁸. Cu catalysed click chemistry was chosen to bioconjugate the enhancer and annealer strands since it has literature precedent for coupling DNA strands with reasonably high yields and it introduces minimal steric bulk into the tripartite sequence. Click chemistry

has since been shown to conjugate two RNA strands together further confirming that click chemistry is compatible with RNA¹⁴⁹.



Scheme 2.2. *Cu catalysed Huisgen [3+2] cycloaddition reaction for tripartite sequence without HEG.*

2.3 – Strategy

The basic design of the tripartite sequence was based on the optimised oligonucleotide sequence identified by Owen et al. Therefore the tripartite sequences will consist of an annealing sequence complementary to positions +2 to +16 of exon 7 in *SMN2* and an artificial ESE sequence derived from three repeats of a known splicing enhancer sequence. The tripartite sequence

however will differ in the way the annealer and enhancer sequences are joined and will include some linker sequences (Scheme 2.3)^{2,140}.

Enhancer and annealer sequences will be prepared by solid phase synthesis including an alkyne on the 5' end of the annealer and an amino modifier on the 3' end of the enhancer to maintain the same directionality as TOES 1 (Scheme 2.3). Both strands will then be deprotected and a post synthesis modification carried out on the enhancer sequence to enable an azide motif to be placed on the 3' end (Scheme 2.3). These strands will then be bioconjugated together using Cu-mediated click chemistry. As previously discussed (*Section 1.7*), having a bioconjugation group in between the annealer and enhancer increases the screening potential for rectifying splicing defects in a more cost effective way. However this system will only work if the bioconjugation group chosen, which is a triazole, is tolerated between the two strands. By adding a flexible linker between the two strands we propose that this will not only increase the conformation landscape to which the enhancer can reach but also negate any detrimental effects introducing a triazole might cause (Scheme 2.3). The flexible linker of choice is a HEG and is incorporated into strands RNA3-RNA11 (Figure 2.19). To determine whether flexibility or length is more beneficial to splice site selection a more rigid linker will also be investigated (Scheme 2.3). The rigid linker called abasic was incorporated into strands RNA14 and RNA15 (Figures 2.17 and 2.30). Comparing the linkers, every three abasics are equivalent to one HEG and the addition of HEG reduces the density of charge in the RNA backbone as only a single charge is present every 1.2 nm relative to ~ 0.4 nm in the abasic case (also approximately equivalent to an RNA base).

These tripartite sequences were tested on a *SMN2* mini gene to determine whether the addition of HEG, abasic and a bioconjugation group effected enhancer activity. *SMN2* system was ideal as exon 7 inclusion levels could be shifted by the addition of these type of sequences.

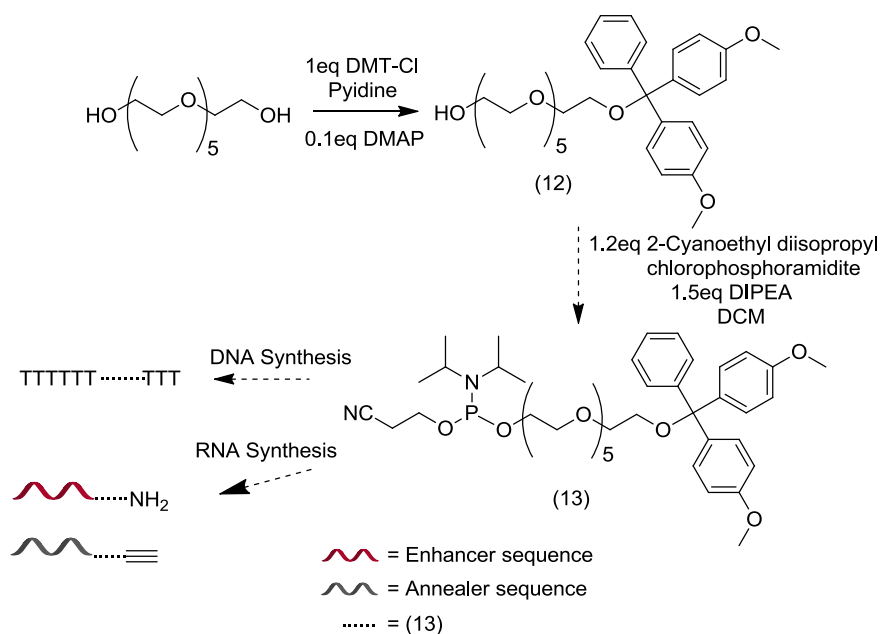
As discussed in *Section 2.1*, TOES 1 (known as GGA) is the best construct to use for the tripartite sequences. Since we are also interested in multivalent systems using Au nanoparticles (*Chapter 4*), pS RNA sequences cannot be used because the pS backbone present in both the annealer and enhancer sequence within TOES 1 could potentially bind to the nanoparticle. Therefore, TOES 27 (known as GGA-O) was chosen as it did not have the pS backbone, but still maintained high levels of exon 7 inclusion of around 5.84 and 1.41 for the *in vivo* and *in vitro* studies, respectively. TOES 27 is a full 2'OMe RNA construct with six repeats of the GGA motif and provides the greatest enhancement of exon 7 inclusion of all the non-pS strands.

2.4 – Results

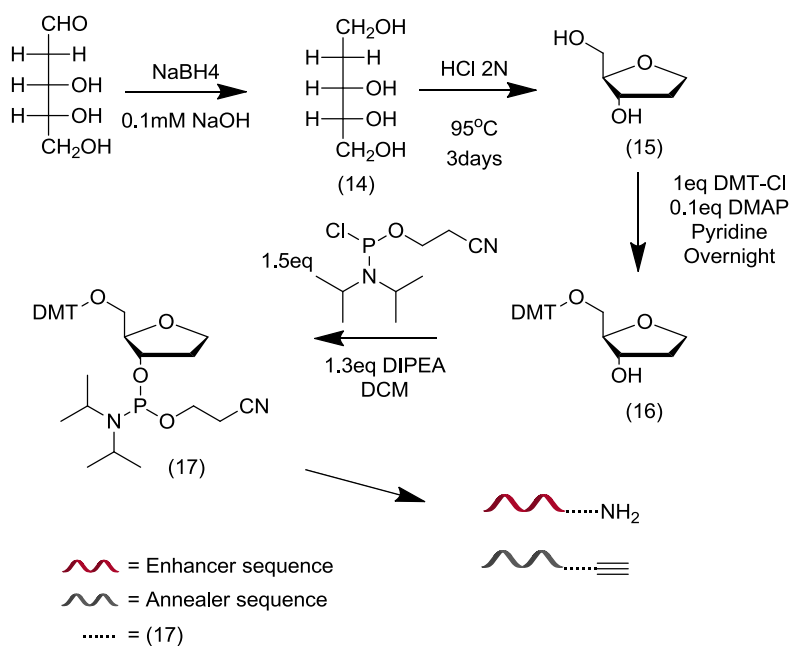
2.4.1 – Linker Synthesis

Two linkers/precursors, compound (**2**) and compound (**8**), were prepared for use in solid phase synthesis with 2'OMe RNA monomers (Scheme 2.3). The linkers were designed to investigate the effects of flexibility on splice site selection. The synthesis of both linkers is shown in Scheme 2.4.

a)



b)



Scheme 2.4. Synthetic route to make linkers for use on a DNA/RNA

synthesiser. a) preparation of compound **13**⁵, b) preparation of compound **17**⁶.

Compound **13** (HEG) was attempted to be synthesised from commercially sourced hexa(ethylene glycol) in two steps. The first step of compound **13** formation was the nucleophilic substitution reaction of the DMT-Cl with

hexa(ethylene glycol) using pyridine as a base (Scheme 2.4) affording compound **12** as a yellow oil in 29% yield. Attachment of the chlorophosphoramidite to compound **12** (Scheme 2.4) provided no desired product after purification. Side products were detected by ¹H NMR which had the peaks for the hexa(ethylene glycol), DMT and the phosphoramidite, however the integrations of these three components did not correspond to one product and further purification of these fractions could not afford separation (Appendix Figure 5.1) . Compound **13** was then commercially sourced from Link Technologies.

Compound **17** was synthesised from 2-deoxy-D-ribose in four steps with compound **16** and **17** isolated (Scheme 2.4 b). Reduction of the aldehyde in 2-deoxy-D-ribose by sodium borohydride afforded compound **14** which was then cyclised using 2N HCl at 95°C for three days affording compound **15**. Crude compound **15** was then DMT protected using DMT-Cl in the presence of pyridine affording compound **16** in 15% yield after column chromatography (neutralised SiO₂). Neutralised silica was used for this reaction as the acidity of the silica could potentially deprotect the DMT protecting group during purification and the silica was re-activated to ensure maximum separation of compound **16** from impurities. Reaction of compound **16** with chlorophosphoramidite afforded compound **17**. Compound **17** was purified by neutralised re-activated silica, using a positive pressure of argon affording compound **17** in 74% yield (Appendix Figure 5.2). Neutralised silica was again used to ensure that no DMT deprotection during chromatography occurred. The column was run under a positive pressure of argon due to the air and moisture sensitive nature of compound **17**. To further minimise exposure of compound **17**

to air and moisture fractions were collected in argon-filled 50ml round bottom flasks and solvent removed directly on a high vacuum line.

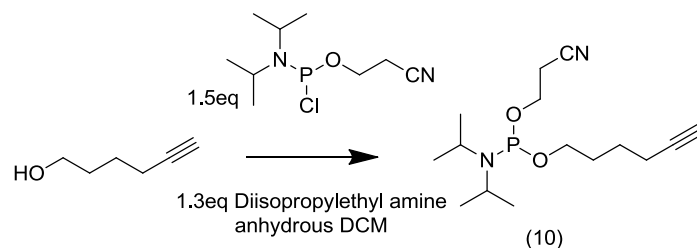
2.4.2 – Preparation of Reactive Groups for Bioconjugation

Scheme 2.5 shows synthetic routes to both compounds **10** and **11**, inserting the bioconjugation groups onto the synthetic 2'OMe RNA strands required for Cu-catalysed Huisgen [3+2] cycloaddition reaction (Scheme 2.3)^{6,150}. Compound **10** was incorporated on the 5' end of the DNA or RNA strand during solid phase synthesis (Scheme 2.3). Compound **11** however requires post synthesis modification of a 3' amino modified DNA/RNA strand (Scheme 2.3).

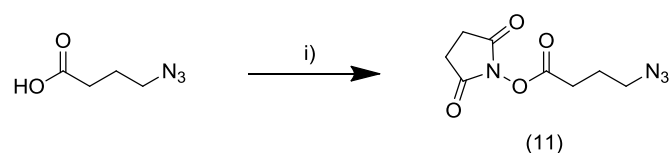
The air-sensitive phosphoramidite **10** was synthesised in one step from hexyn-1-ol in 83 % yield (Appendix Figure 5.3). The reaction was carried out under anhydrous conditions with the product purified by neutralised re-activated silica under a positive pressure of argon. Five 50ml fractions were collected from the column and the solvent was removed on a high vacuum line to reduce the contact with water vapour and air.

The NHS ester **11** was synthesised in one step from 4-azidobutanoic acid and N-hydroxysuccinimide using HOBt in anhydrous dichloromethane (DCM). Anhydrous DCM was used for the reaction to reduces hydrolysis of the NHS ester and the product was purified by normal silica chromatography. Compound **11** was obtained from the column as a white solid in 34% yield (Appendix Figure 5.4).

a)



b)



Scheme 2.5. Synthetic route to provide reactive intermediates for Cu-catalysed Huisgen [3+2] cycloaddition. a) preparation of compound **10** and b) preparation of compound **11**. i) 1.5eq *N*-Hydroxysuccinimid, 1.5eq HOBt, 1.5eq EDC and anhydrous DCM.

2.4.3 – Synthesis of Enhancer and Annealer Strands for Bioconjugation

2.4.3.1 – 2'OMe RNA Solid Phase Synthesis

As efficient incorporation was obtained using compound **2**, 2'OMe RNA versions of enhancer and annealer strands were made (Table 2.3).

Table 2.3 highlights all the enhancer and annealer strands synthesised for the preparation of the two halves of the TOES.

Table 2.3. All deconstructed TOES sequences for solid phase synthesis. Where *X* = HEG, *Z*= C6 amino modifier and *Y* = compound **10**

Strand	Enhancer Sequence – 5' to 3'	Annealer Sequence – 5' to 3'
RNA1_A	A ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ A ₀ Z	-
RNA1_C	A ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ A ₀ XZ	-
RNA1_E	A ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ A ₀ XXXXXXXXXXZ	-
RNA1_G	A ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ G ₀ G ₀ A ₀ G ₀ G ₀ A ₀ C ₀ A ₀ XXXZ	-
RNA2_A	-	YG ₀ A ₀ U ₀ U ₀ U ₀ G ₀ U ₀ C ₀ U ₀ A ₀ A ₀ A ₀ C ₀
RNA2_B	-	YXG ₀ A ₀ U ₀ U ₀ U ₀ G ₀ U ₀ C ₀ U ₀ A ₀ A ₀ A ₀ C ₀
RNA2_C	-	YXXXXXXXXXXG ₀ A ₀ U ₀ U ₀ U ₀ G ₀ U ₀ C ₀ U ₀ A ₀ A ₀ A ₀ C ₀

The 2'OMe RNA bases have extra bulk at the 2' position of the ribose ring, so coupling times for stands RNA2_A, RNA2_B, RNA2_C and RNA1_A were lengthened to 11 minutes compared to 5 minutes for the DNA versions. Coupling times were lengthened further for RNA1_C, RNA1_E and RNA1_G to 18 minutes as these required coupling of 2'OMe RNA bases onto flexible linkers at the 3' end.

RNA2 strands were purified by RP-HPLC using a standard protocol (*Section 2.5.1.2*) (Figure 2.11 a). RNA2 strands showed one major peak which could be successfully purified by RP-HPLC (Figure 2.11 a).

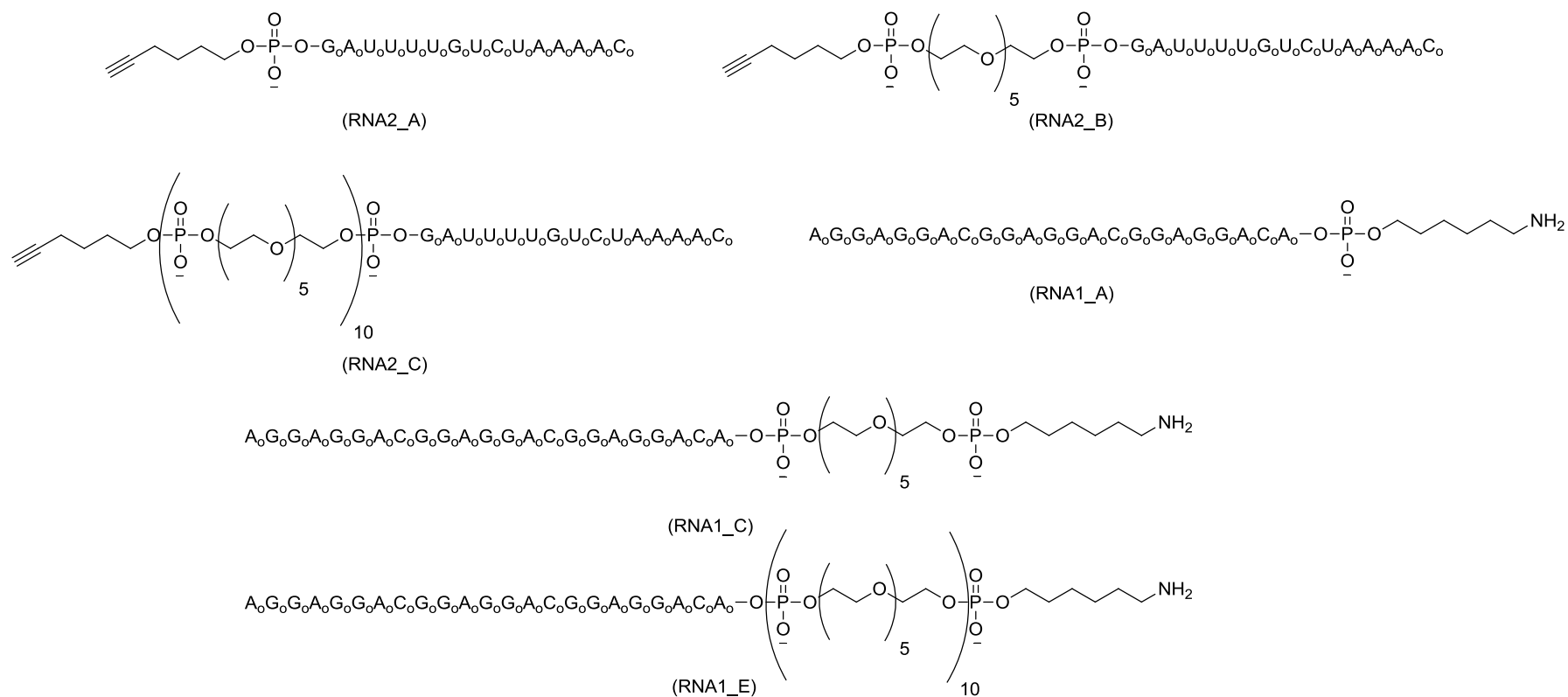


Figure 2.4. Sequences of the RNA1 series and RNA2 series incorporating linker 10. o indicates that bases are 2'OMe RNA.

The purification of the RNA1 series however was more problematic as RNA1 could not be successfully purified by RP-HPLC with RNA1_A showing a multitude of peaks and the presence of a broad peak in the RP-HPLC spectra (Figure 2.11 b). All the peaks isolated from the purification of RNA1_A were checked by MALDI-TOF, but could not be identified (Appendix Figure 5.5).

It was later discovered that purification of RNA1_A by RP-HPLC was not possible due to the nature of the RNA1 sequence (Section 2.4.14) which caused all the RNA1_A to be lost during RP-HPLC. Ion exchange purification was carried out on RNA1_C and RNA1_E as it is a good purification technique for sequences with secondary structures. RNA1_C was successfully purified by running a gradient from 20mM to 1M NaCl with 10mM Hepes over 40 minutes of elution. RNA1_E was also successfully purified using this method, but an extended time of 80 minutes was required for efficient separation (Figure 2.5).

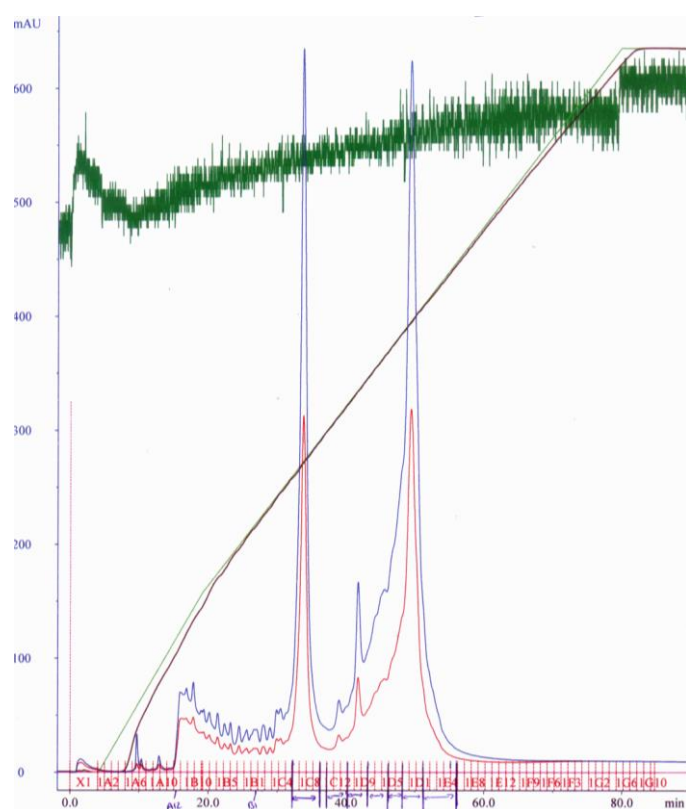


Figure 2.5. Ion exchange trace for RNA1_E purification running a gradient from 0% to 100% over 80 minutes. Where buffer A is 20mM NaCl with 10mM Hepes and buffer B is 1M NaCl with 10mM Hepes. The blue line indicates absorption at 260nm, the red line absorption at 280nm and the light green line the

gradient. RNA1_C trace shown in appendix (Figure 5.7).

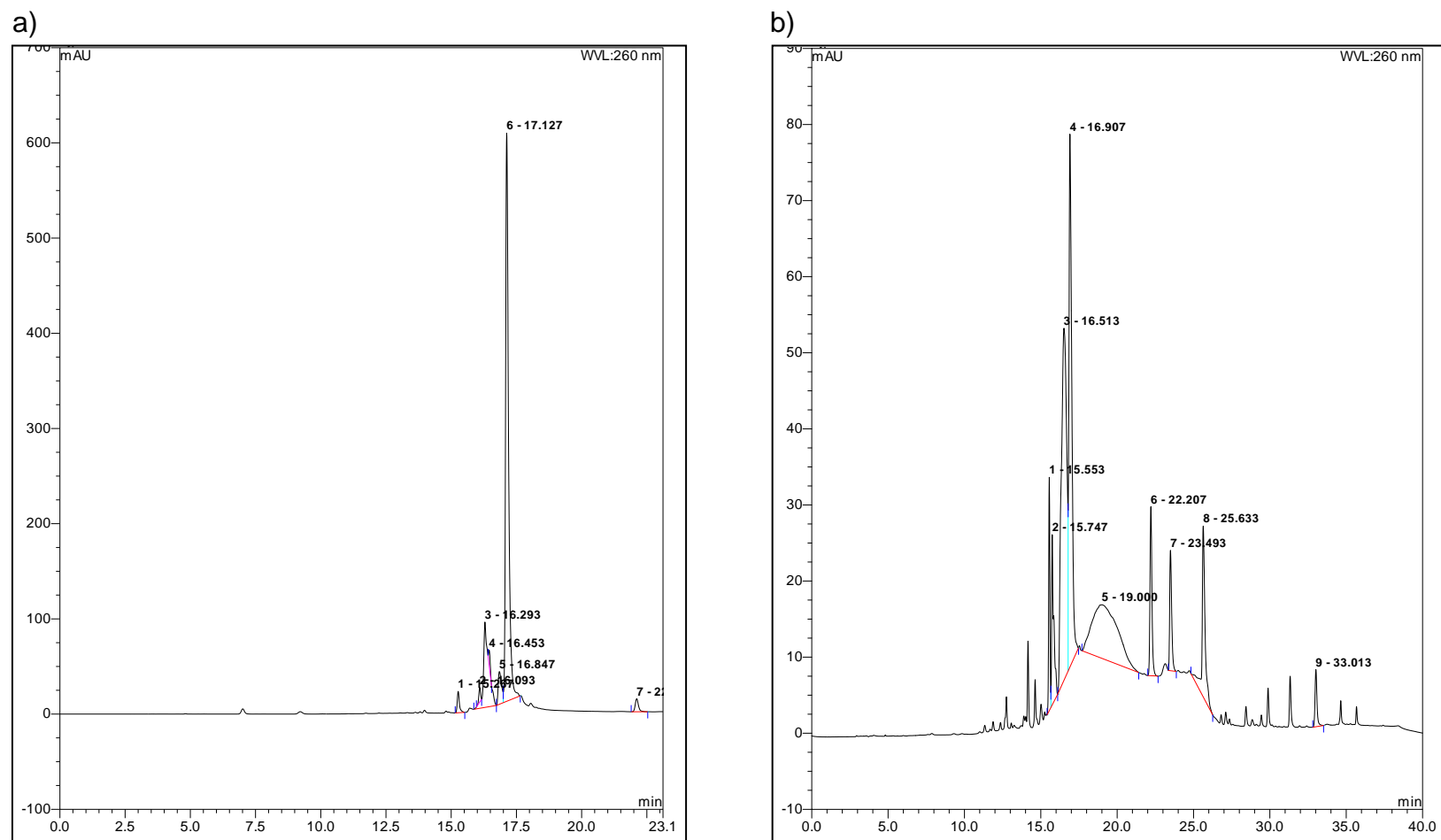


Figure 2.6. RP-HPLC traces for a) RNA2_A and b) RNA1_A. RP-HPLC was carried out using a standard protocol with 5% - 60% gradients (Section 2.5.1.2). RNA2_B and RNA2_C RP-HPLC traces shown in appendix (Figures 5.6).

The ion exchange traces for RNA1_C and RNA1_E both showed two main peaks (Figure 2.5). To determine which peaks corresponded to the desired products and whether the fractions were pure, 10% denaturing polyacrylamide gels were used (Figure 2.7).

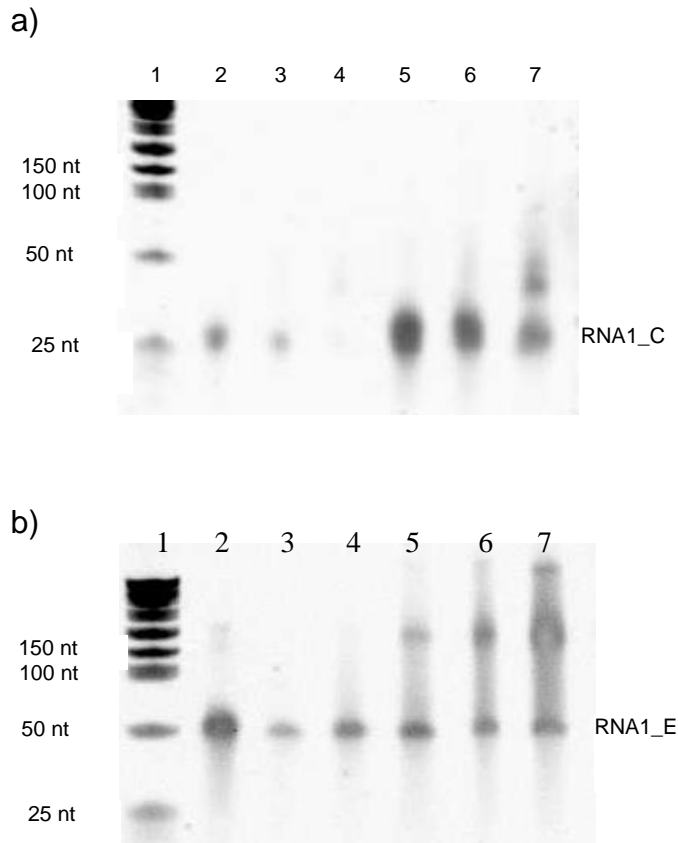


Figure 2.7. 10% denaturing polyacrylamide gel of ion exchange column fractions. a) RNA1_C fraction. Lane 1) DNA ladder. Lanes 2-7) Fractions 1-6. b) RNA1_E fraction. Lane 1) DNA ladder. Lanes 2-7) Fractions 1-6.

Fractions 1-4 were combined for RNA1_C and fractions 1-3 combined for RNA1_E as these fractions appeared to be pure on the polyacrylamide gel. These combined fractions were analysed by polyacrylamide gel electrophoresis (PAGE) along with RP-HPLC purified RNA2_A, B, C and re-prepared RNA1_A (Figure 2.8).

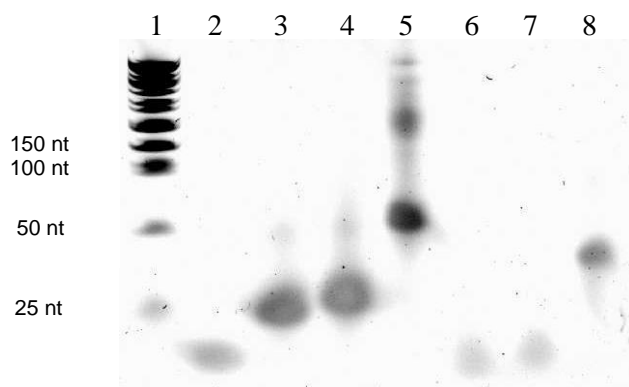


Figure 2.8. 10% denaturing polyacrylamide gel. Lane 1) DNA marker. Lane 2) RNA 2. Lane 3) Crude RNA1_A. Lane 4) RNA1_C fractions 1-4 combined. Lane 5) RNA1_E fractions 1-3 combined. Lane 6) RP-HPLC purified RNA2_A. Lane 7) RP-HPLC purified RNA2_B. Lane 8) RP-HPLC purified RNA2_C.

The gel revealed the RNA2 series was successfully purified by RP-HPLC as there was only a single band present in lanes 2, 6, 7 and 8. More than one band was observed in both the RNA1_C and E combined fraction samples in lanes 4 and 5 which had not been seen in Figure 9 for these fractions before they were combined. It was thought that these extra bands may be due to secondary structure formation¹⁵¹. This was also thought to be the reason for the extra band seen in lane 3 for the RNA1_A sample. The presence of a secondary structure in RNA1_A was investigated using UV-visible isothermal melting (Figure 2.9). RNA1_A was dissolved in a buffer of KCl 100mM, potassium phosphate (KP) 10mM, pH 7 in water giving a final RNA1_A concentration of 4 μ M.

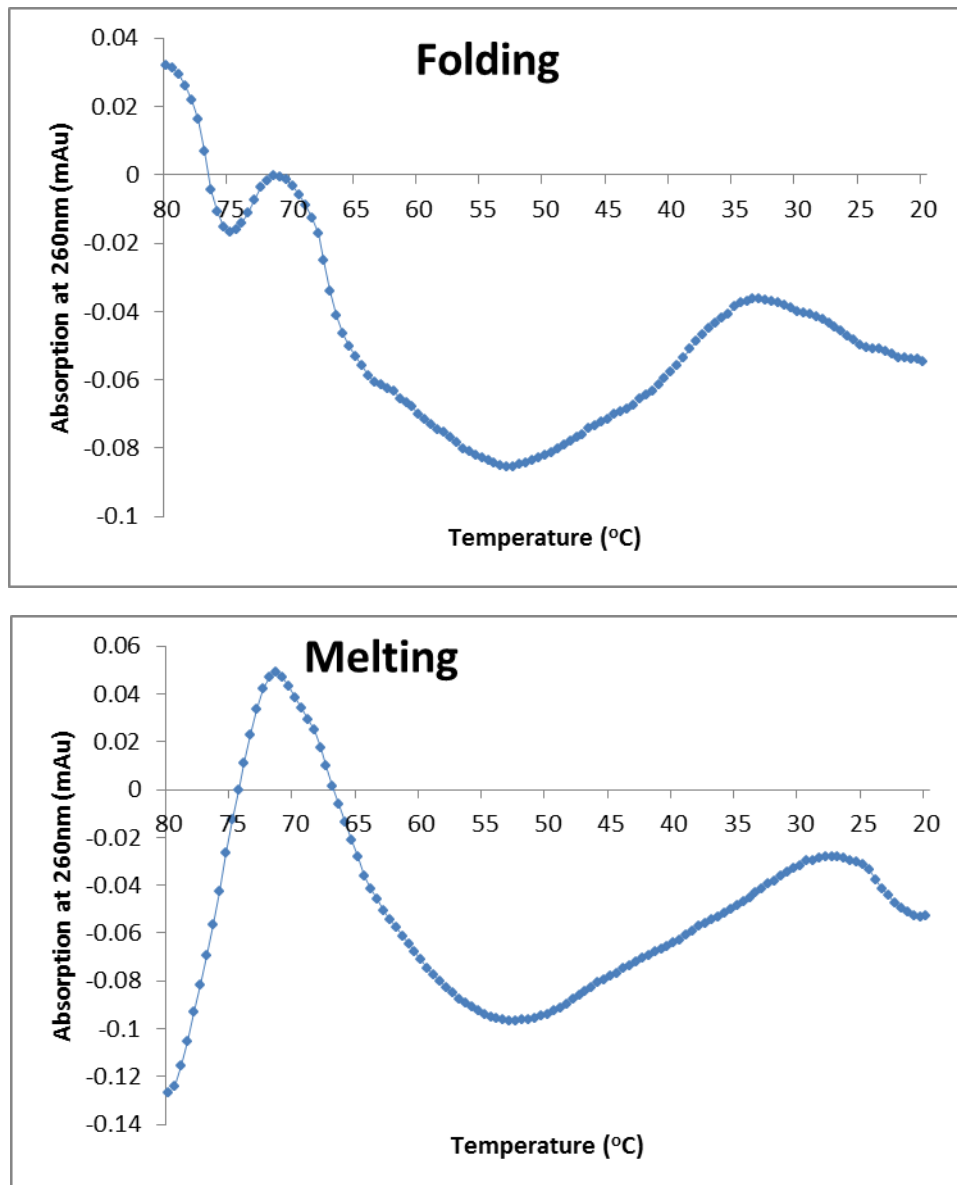


Figure 2.9. T_m folding and melting graphs for RNA1_A samples dissolved in KCl 100mM and KP 10mM at pH~7. T_m measured, for ramp rates of 0.2°C per minute, on both folding and melting runs.

The T_m of RNA1_A was found to be ~54°C, indicating the presence of a secondary structure. This, along with the fact that secondary structures shift mobility of strands in PAGE strongly suggests that the top bands on Figure 2.8 in lanes 3-5 are secondary structures of the RNA sequences. This was further confirmed by heating RNA1_C and RNA1_E to 80°C prior to PAGE, which resulted in loss of the higher bands as the heat-treatment denatured the

secondary structure (Figure 2.10). Secondary structure formation also accounts for the extra peaks seen on the RP-HPLC trace for RNA1_A (Figure 2.6 b) and also for the extra bands in fractions 4-6 after ion exchange of RNA1_E (Figure 2.7). The RNA1 strands showed no peaks or bands for anything other than the product and the secondary structure when analysed by ion exchange and gel electrophoresis. Therefore crude RNA1 sequences only require desalting through an Illustra NAPTM 25 column to purify them. Further investigations into the secondary structure of RNA1 are discussed in *Section 2.4.14*.

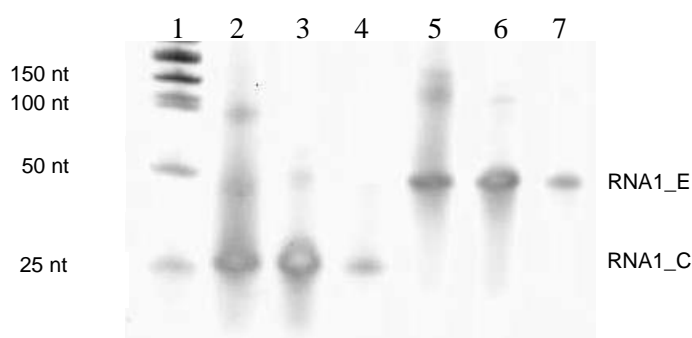


Figure 2.10. 10% denaturing polyacrylamide gel of RNA1_C and RNA1_E samples. Lane 1) DNA ladder. Lane 2) RNA1_C. Lane 3) RNA1_C heat treated at 80°C. Lane 4) 1 in 100 dilution of RNA1_C. Lane 5) RNA1_E. Lane 6) RNA1_E heat treated at 80°C. Lane 7) 1 in 100 dilution of RNA1_E.

As the RNA1 series were found to be pure, they were analysed by MALDI-TOF, along with the RNA2 series, to confirm strand sequence. RNA1_A, RNA1_C, RNA1_E, RNA2_A, RNA2_B and RNA2_C were all identified successfully by MALDI-TOF (Table 2.4).

Table 2.4. RNA sequences prepared by solid phase synthesis. MALDI and mass spec data included in appendix (Figures 5.8)

Sequence	Expected MW	MW detected [M+H]	% Yield after purification
RNA1_A	8131.29	8131.83	8%
RNA1_C	8475.58	8476.28	16%
RNA1_E	11574.15	11575.21	15%
RNA2_A	5092.23	5092.82	20%
RNA2_B	5436.51	5436.76	24%
RNA2_C	8535.08	8535.52	21%
RNA1_G	8670.29	8670.41	-

Findings from splicing experiments (*Section 3.5.1.2*) revealed the best tripartite sequence for increasing exon 7 inclusion levels was RNA6. Abasic versions of RNA6 were constructed to investigate how important flexibility was in the tripartite system. These abasic versions were named RNA1_G and RNA14 (Figure 2.12). RNA1_G and RNA14 were synthesised by solid phase synthesis with RNA14 was purified by PAGE and RNA1_G purified by an Illustra NAPTM 25 column. Both products were further checked by 20% denaturing PAGE and RNA1_G was analysed by mass spec (Figure 2.11 and Table 2.3).

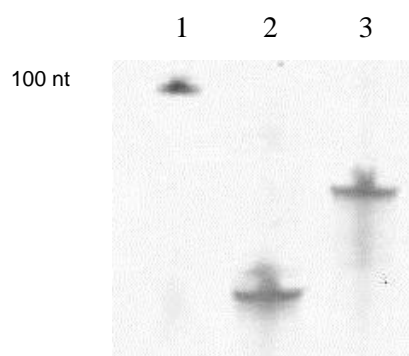


Figure 2.11. 20% denaturing polyacrylamide gel of RNA1_G and RNA14 samples. Lane 1) DNA ladder. Lane 2) purified RNA1_G. Lane 3) RNA14 respectively.

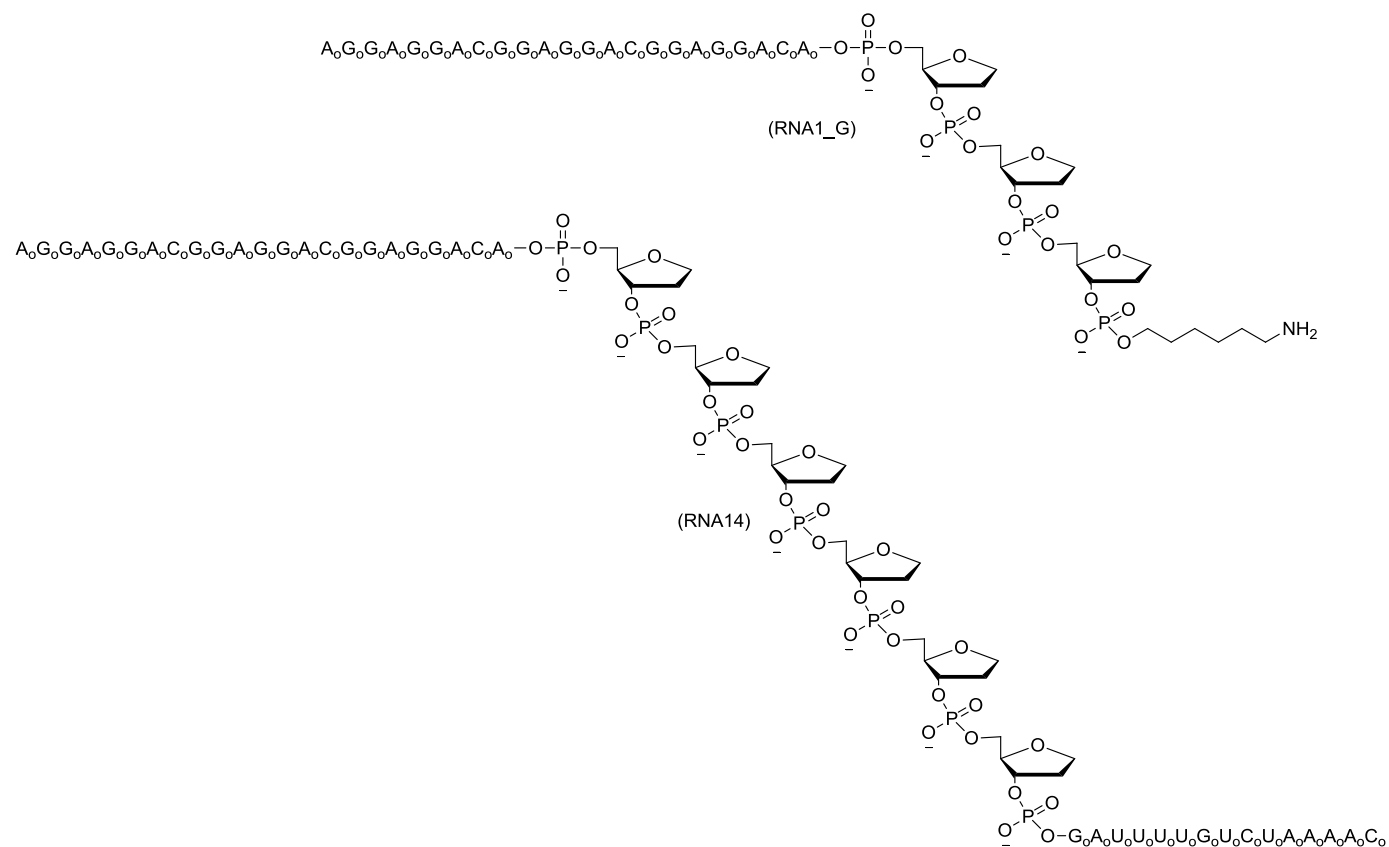


Figure 2.12. Sequences RNA1_G and RNA14 incorporating linkers 17. _o indicates that bases are 2'OMe RNA.

2.4.3.2 – Modification of the 3' Amino Modified Enhancers by NHS Azide Conjugation (Compound 11)

For RNA1_A, RNA1_C, RNA1_E, and RNA1_G strands to undergo a Cu-catalysed Huisgen [3+2] cycloaddition reaction an azido moiety needed to be attached (Scheme 2.3). To attach the azide motif onto the RNA strands, a post synthesis modification was undertaken with compound 4 to give RNA1_B, RNA1_D, RNA1_F. and RNA1_H (Figure 2.13) which were characterised by mass spec (Table 2.5). Reactions were carried out in a 7:3 mix of DNA grade acetonitrile/aqueous sodium hydrogen carbonate (0.025M, pH~8) to stabilise compound **11** during the reaction. Since the azide moiety did not affect the mobility shift of the strands on a denaturing polyacrylamide gel, the products were worked up, the acetonitrile removed and the products run through Illustra NAP[™] 25 columns to remove unconjugated compound **11** and hydrolysed product.

Table 2.5. Sequences made by conjugation to compound **11** from RNA1_A, RNA1_C, RNA1_E and RNA1_G. Mass spec data included in appendix (Figures 5.9).

Sequence	Expected MW	MW detected [M+H]
RNA1_B	8242.39	8243.63
RNA1_D	8586.68	8586.68
RNA1_F	11685.25	11683.65
RNA1_H	8782.31	8781.53

These strands were then ready for click chemistry reactions.

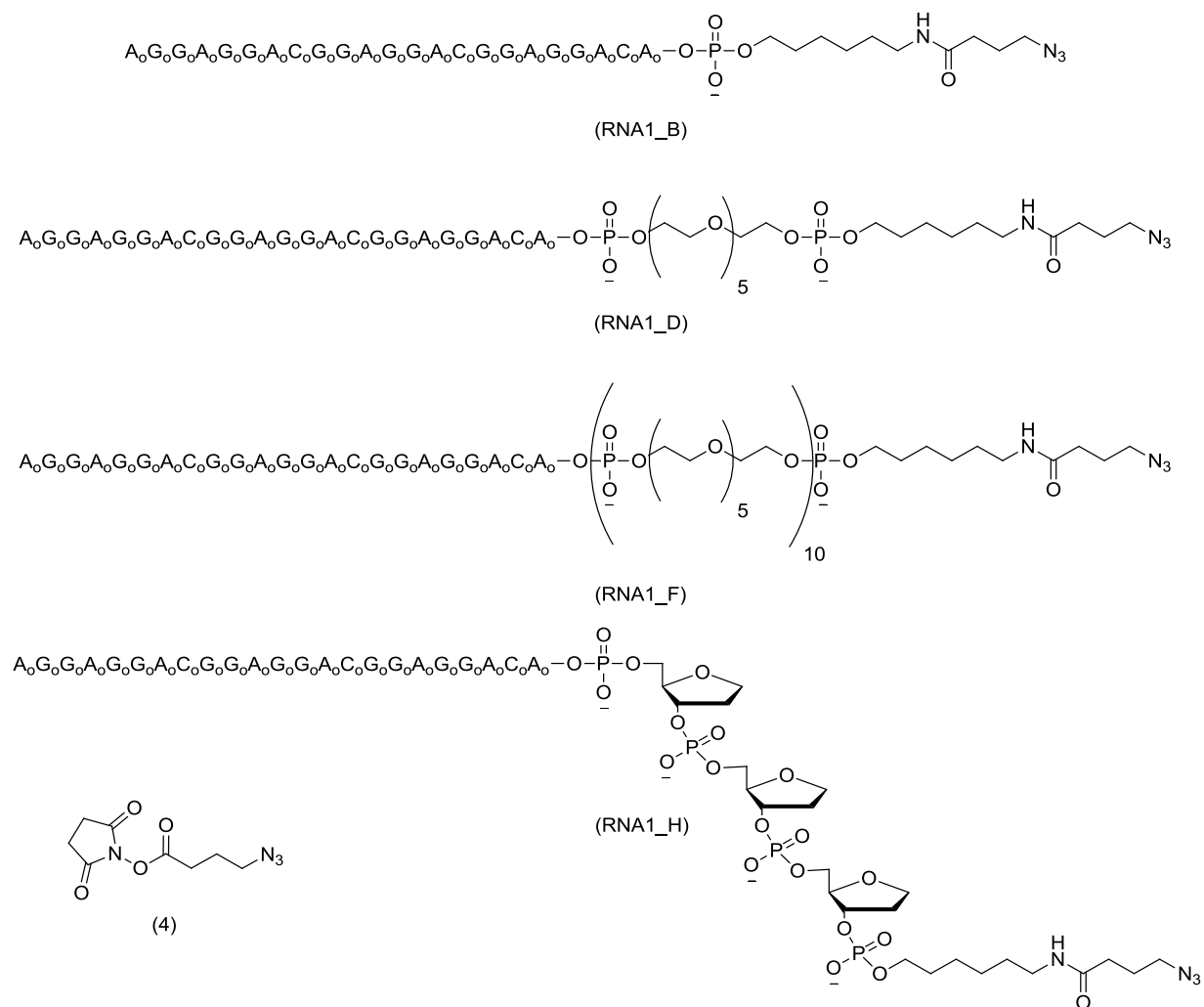


Figure 2.13. RNA1_A, RNA1_C, RNA1_E and RNA1_G products after NHS reaction with compound 11.

2.4.4 – Click Chemistry Reactions

2.4.4.1 – Trial Click Chemistry

To investigate whether click chemistry was possible between two 2'OMe RNA strands, products RNA3-RNA6 and RNA11 were bioconjugated (Figure 2.15).

RNA6 was prepared to test whether the azide moiety had been attached to the RNA1_D. Two amounts of copper were trialled, 100eq and 10eq. These concentrations were used as 100eq had been shown to bioconjugate two DNA strands and 10eq of Cu was chosen as high concentrations of Cu can cause problems with DNA and RNA (Section 1.7.1.3)¹⁴⁸. A time course was taken to investigate how long the reaction would take with samples being desalted and then separated by 10% denaturing PAGE (Figure 2.14).

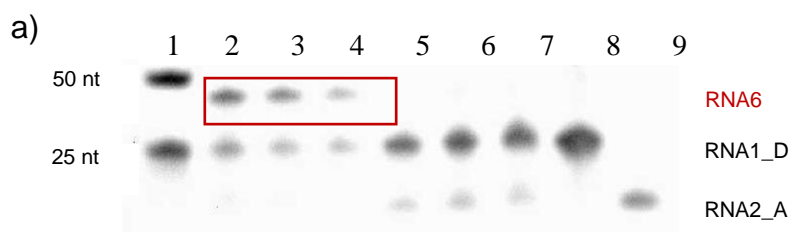


Figure 2.14. 10% denaturing polyacrylamide gels of reaction samples from bioconjugation of RNA1_D and RNA2_A. **a)** Lane 1) DNA marker. Lanes 2-4) Click reaction of RNA1_D and RNA2_A with 100eq of CuTBTA at 15, 30 and 60 minute, respectively. Lanes 5-7) Click reaction of RNA1_D and RNA2_A with 10eq of CuTBTA at 15, 30 and 60 minute, respectively. Lane 8) RNA1_D. Lane 9) RNA2_A. Products shown in red box.

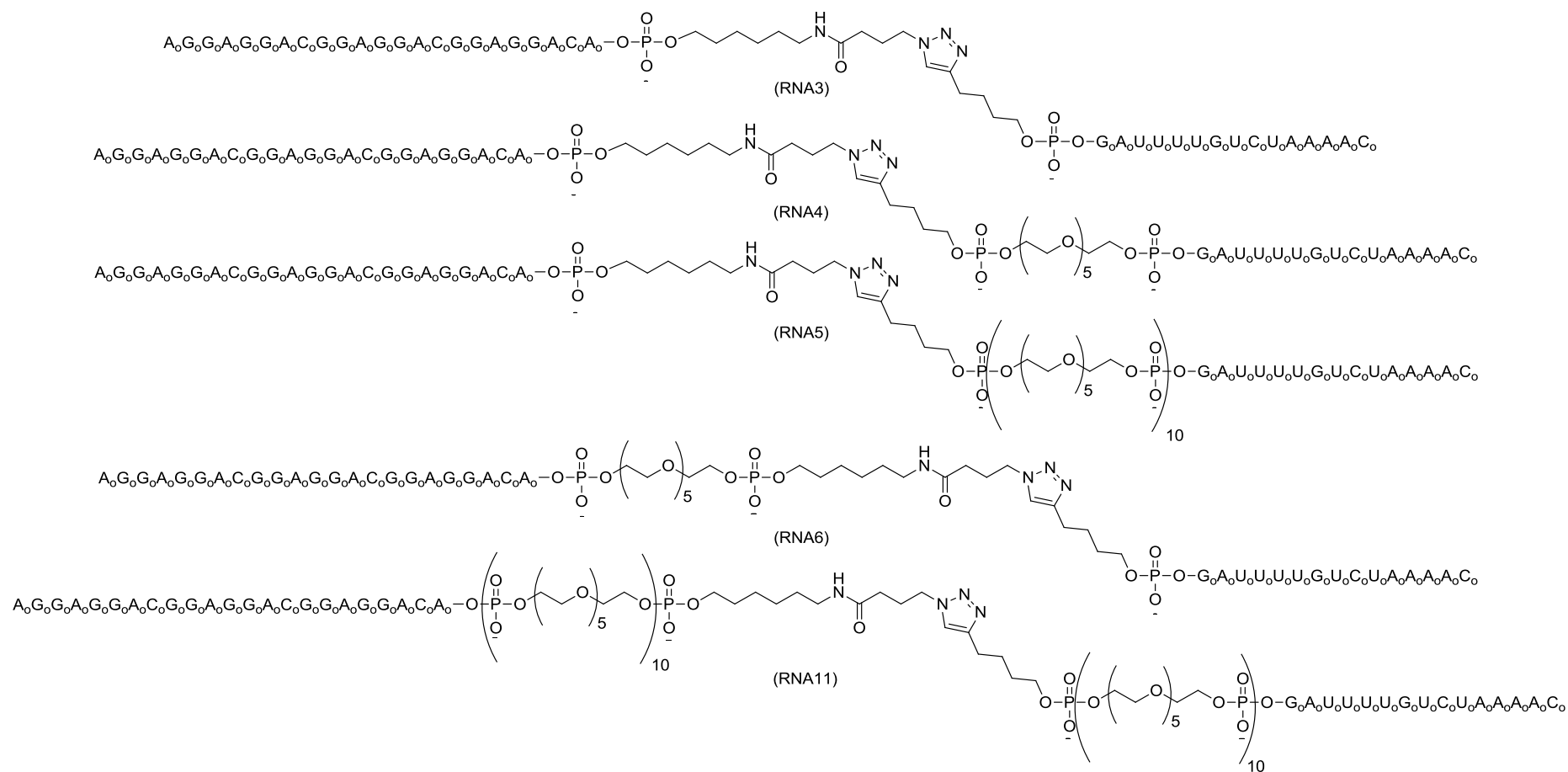


Figure 2.15. Products of Cu-catalysed Huisgen [3+2] cycloaddition reactions characterised by 10% PAGE.

Reactions with 10eq of CuTBTA showed no product formation, however reactions with 100eq lead to the disappearance of RNA2_A and the appearance of a new band. The new band was the correct size for RNA6 and had a different electrophoretic mobility than the starting materials and the controls. No RNA2_A was seen on gel) in lanes 2-4 where 100eq of CuTBTA were used, but the RNA1_D band remained visible. Since 1eq of both RNA1_D and RNA2_A were added to the reaction and since lanes 5-8 showed only faint RNA2_A bands when no reaction occurred, it was thought that the SYBR gold staining of RNA2_A may have been inefficient.

As the click chemistry reaction between RNA1_D and RNA2_A was successful, click chemistry was trialled for the longest product, RNA11. RNA11 was conjugated from RNA1_F and RNA2_C with both reagents being added at 1eq. The CuTBTA level was fixed at 100eq and the reaction carried out as above. Samples were desalted followed by separation by 10% denaturing PAGE (Figure 2.16).



Figure 2.16. 10% denaturing polyacrylamide gel of reaction samples from bioconjugation of RNA1_F and RNA2_C. Lane 1) DNA marker. Lanes 2-5) Reaction at 0, 15, 30 and 60 minute time points, respectively. Lanes 6-7) RNA1_F with and without Cu. Lanes 8-9) RNA2_C with and without Cu. Products shown in red box.

A new band was seen in lanes 3-5 which corresponded to the expected size of RNA11. Both starting materials, RNA1_F and RNA2_C, were still present even after the 1 hour time point, however the reaction had stopped after the 15 minute time point with ~50% product conversion. In an attempt to force the reaction to go further more equivalents of RNA2 strands were added to the reaction as the RNA2 series were easier to synthesise and purify.

2.4.5 – All 2'OMe RNA Click Chemistry Reactions

To check whether all 2'OMe RNA click chemistry reactions were possible, bioconjugation reactions were carried out as indicated in Table 2.6. The structures of the products of are shown in Figure 2.17.

Table 2.6. *Reactants and products for all 2'OMe RNA click chemistry bioconjugation reactions undertaken*

Azide starting material	Alkyne starting material	Product produced
RNA1_B	RNA2_A	RNA3
RNA1_B	RNA2_B	RNA4
RNA1_B	RNA2_C	RNA5
RNA1_D	RNA2_A	RNA6
RNA1_D	RNA2_B	RNA7
RNA1_D	RNA2_C	RNA8
RNA1_F	RNA2_A	RNA9
RNA1_F	RNA2_B	RNA10
RNA1_F	RNA2_C	RNA11
RNA1_H	RNA2_A	RNA15

The reactions that produced RNA3-11 were desalted before being analysed by 10% denaturing PAGE (Figure 2.18). The reaction producing RNA 15 was carried straight through to click chemistry scale up.

In all the reaction samples, a band was observed which had a different electrophoretic mobility to the starting materials and which corresponded to the approximate size of the desired products (Figure 2.18). PAGE revealed that in all reactions there was no alkyne starting materials (RNA2 series) remaining after 15 minutes, but this was likely to be due to poor SYBR gold staining of the RNA2 series. Approximate percentage conversions for the reactions can be seen in Table 2.7.

Table 2.7. *Approximate percentage product conversions for click chemistry reactions*

Product Name	Starting Materials	Approximate Percentage Conversion
RNA3	RNA1_B RNA2_A	46%
RNA4	RNA1_B RNA2_B	53%
RNA5	RNA1_B RNA2_C	50%
RNA6	RNA1_D RNA2_A	55%
RNA7	RNA1_D RNA2_B	49%
RNA8	RNA1_D RNA2_C	50%
RNA9	RNA1_F RNA2_A	36%
RNA10	RNA1_F RNA2_B	39%
RNA11	RNA1_F RNA2_C	42%

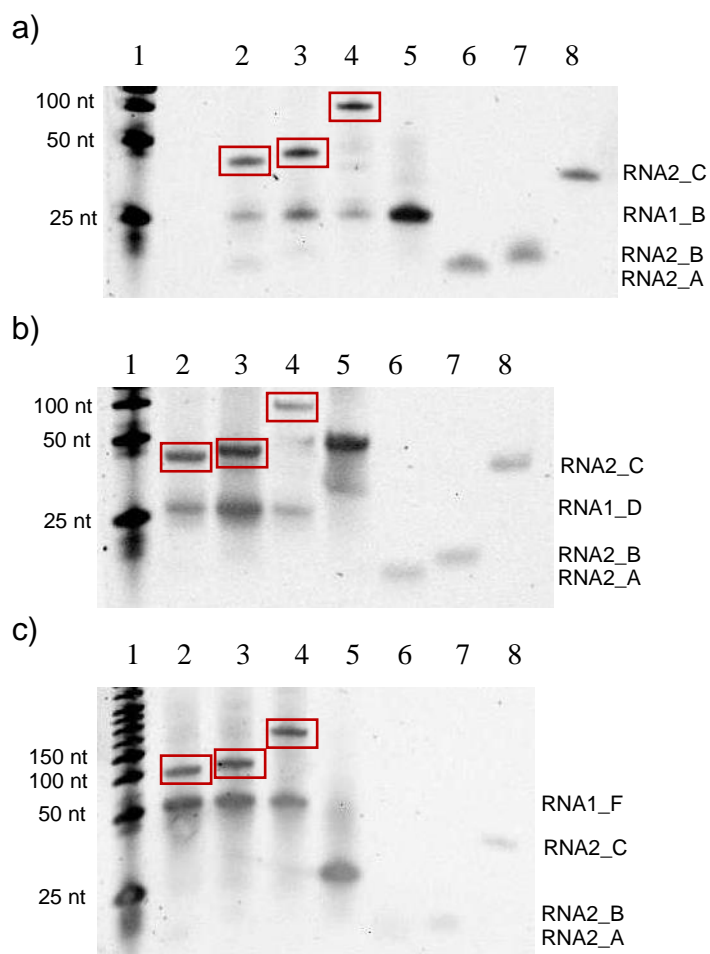


Figure 2.18. 10% denaturing polyacrylamide gels of 15 minute time point samples from click chemistry reactions. **a)** Lane 1) DNA ladder. Lane 2) Reaction between RNA1_B and RNA2_A. Lane 3) Reaction between RNA1_B and RNA2_B. Lane 4) Reaction between RNA1_B and RNA2_C. Lane 5-8) Controls: RNA1_B, RNA2_A, RNA2_B and RNA2_C, respectively. **b)** Lane 1) DNA ladder. Lane 2) Reaction between RNA1_D and RNA2_A. Lane 3) Reaction between RNA1_D and RNA2_B. Lane 4) Reaction between RNA1_D and RNA2_C. Lane 5-8) Controls: RNA1_F, RNA2_A, RNA2_B and RNA2_C, respectively. **c)** Lane 1) DNA ladder. Lane 2) Reaction between RNA1_F and RNA2_A. Lane 3) Reaction between RNA1_F and RNA2_B. Lane 4) Reaction between RNA1_F and RNA2_C. Lane 5-8) Controls: RNA1_D, RNA2_A, RNA2_B and RNA2_C, respectively. Products shown in red boxes.

2.4.6 - 2'OMe RNA Click Chemistry Scale Up

As all the reactions in Table 2.6 achieved a yield of greater than 35%, scale up reactions of RNA3 to RNA11 and RNA15 were carried out using 1eq of both azide and alkyne and 100eq of CuTBTA. Twelve side by side reactions for each tripartite sequence were carried out to bioconjugate enough material for the splicing experiments. Reactions were run for 15 minutes since all previous click chemistry reactions undertaken had showed no more product formation after 15 minutes.

A batch of RNA6 was split into three and used to investigate purification conditions. Ion exchange and gel electrophoresis were trialled as HPLC purifications were not possible due to the presence of the RNA1 strand. Gel electrophoresis purification gave a yield of 3% compared to the 0.3% yield obtained from ion exchange purification. The reason the ion exchange purification was less efficient than the gel electrophoresis was due to the small reaction scale and inefficiency of product elution from the column.

Polyacrylamide gels were therefore used to purify the tripartite sequences. Mini gels were used because they are smaller and thicker than the larger gels which make post staining, gel manipulation and band excision easier. The gels were stained with SYBR[™] gold for 20 minutes then imaged using the VersaDoc[™] imager. The product bands were excised and worked up as described in (Section 2.5.1.5). As with the previous experiments the click chemistry reactions never reached completion, therefore low yields were obtained for the bioconjugation reactions (Table 2.8). The starting materials were also isolated from the reactions to re-use to minimise losses during the reactions. The

products were analysed against starting materials by 10% denaturing PAGE to check their purity (Figure 2.19).

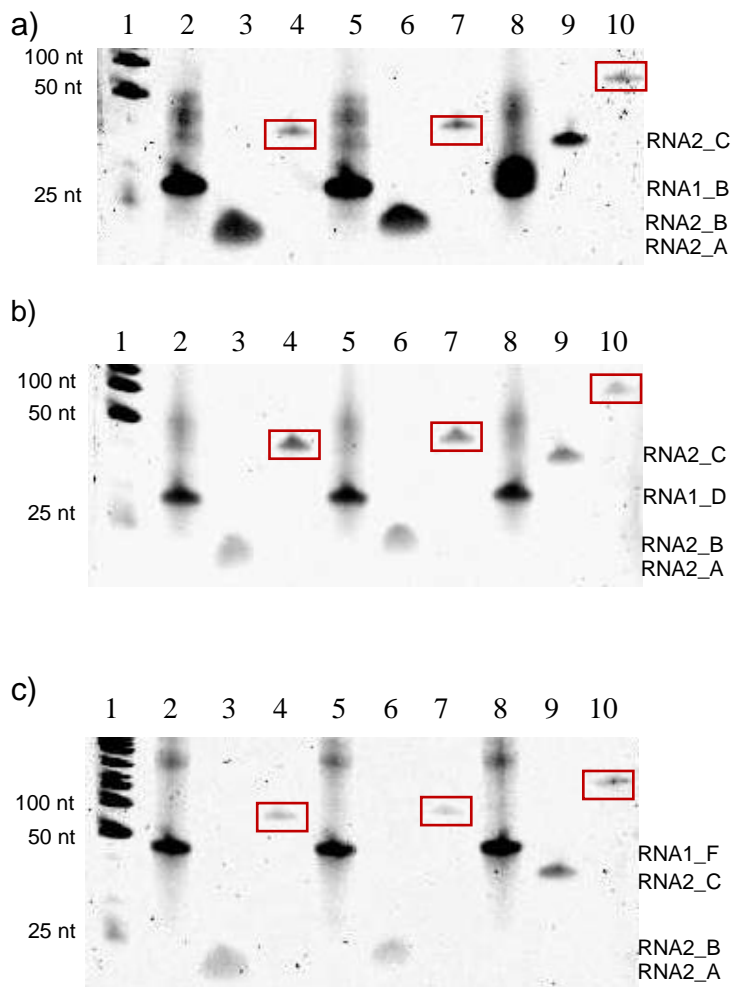


Figure 2.19. 10% denaturing polyacrylamide gels of purified RNA3-RNA11 and their starting materials. **a)** Lane 1) DNA ladder. Lanes 2, 5 & 8) RNA1_B. Lane 3) RNA2_A. Lane 4) RNA3. Lane 6) RNA2_B. Lane 7) RNA4. Lane 9) RNA2_C. Lane 10) RNA5. **b)** Lane 1) DNA ladder. Lanes 2, 5 & 8) RNA1_D. Lane 3) RNA2_A. Lane 4) RNA6. Lane 6) RNA2_B. Lane 7) RNA7. Lane 9) RNA2_C. Lane 10) RNA8. **c)** Lane 1) DNA ladder. Lanes 2, 5 & 8) RNA1_F. Lane 3) RNA2_A. Lane 4) RNA9. Lane 6) RNA2_B. Lane 7) RNA10. Lane 9) RNA2_C. Lane 10) RNA11. Products shown in red boxes.

Table 2.8. *Percentage isolated yields for products RNA3-RNA11 and RNA15 and their starting materials*

Product Name	Starting Materials	Isolated Yields
RNA3	RNA1_B RNA2_A	9%
RNA4	RNA1_B RNA2_B	15%
RNA5	RNA1_B RNA2_C	7%
RNA6	RNA1_D RNA2_A	19%
RNA7	RNA1_D RNA2_B	15%
RNA8	RNA1_D RNA2_C	14%
RNA9	RNA1_F RNA2_A	10%
RNA10	RNA1_F RNA2_B	8%
RNA11	RNA1_F RNA2_C	5%
RNA15	RNA_H RNA2_A	12%

2.4.7 – Phosphothioate Synthesis

Phosphothioate (pS) backbones have been shown to increase protein binding². Therefore pS 2'OMe RNA versions of tripartite sequences were synthesised to determine whether pS would increase the activity of the enhancer sequence and promote greater levels of exon 7 inclusion in *SMN2*. As phosphothioate click chemistry was not documented at the time work was carried out, DNA versions were made and trialled before preparing 2'OMe RNA versions. pS DNA3 was made using an “aged” PADS oxidiser (Figure 2.21). Deprotection was achieved by 1:1 triethylamine/acetonitrile prior to cleavage using

35% ammonium hydroxide. MALDI-TOF of pS DNA confirmed the presence of a full phosphothioate strand (Figure 2.20) which meant that the 2'OMe RNA versions could be made.

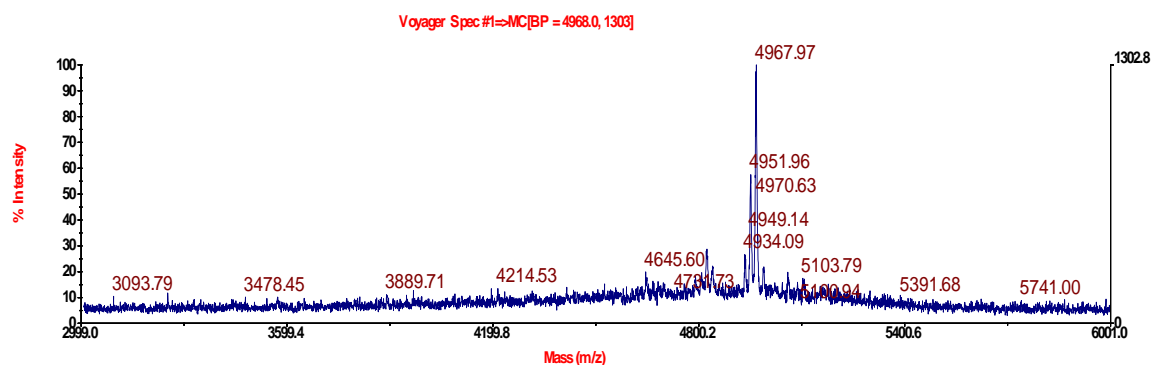


Figure 2.20. MALDI-TOF data run in negative mode for pS DNA3 where the expected phosphothioate $[M+H] = 4965.05$.

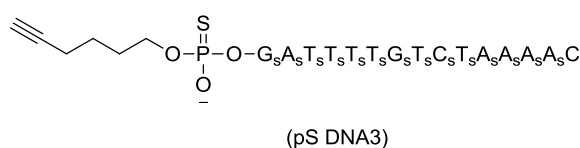


Figure 2.21. Structures of pS DNA3 phosphothioate strands synthesised on a ABI 394 DNA/RNA synthesiser.

2.4.8 – pS 2'OMe RNA Solid Phase Synthesis

Phosphothioate versions of the RNA1 and RNA2 series were synthesised to discover whether they would increase protein binding and therefore tripartite sequence activity (Figure 2.23).

These sequences were synthesised and worked up as described in Section 2.4.7 with all the strands being treated with the triethylamine/acetonitrile mix.

After several attempts at MALDI-TOF using different matrices, voltages, laser powers and sample concentrations could determine masses for these compounds (Appendix Figures 5.12). Therefore the sequences were used crude in the click chemistry reactions and analysed by PAGE to determine whether the strands had the correct electrophoretic mobility (Figure 2.25). Before pS RNA1_A, pS RNA1_C and pS RNA1_E were used in click chemistry reactions, their amino terminal groups were reacted with compound **11** to give pS RNA1_B, pS RNA1_D and ps RNA1_F, respectively (Figure 2.22, *Section 2.4.7.1*).

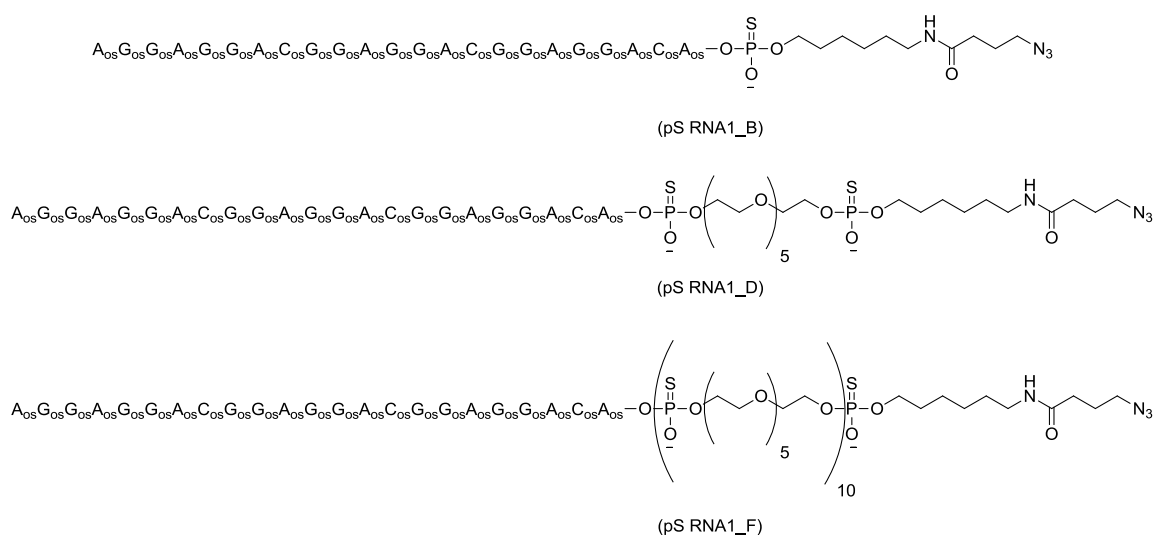


Figure 2.22. Products formed after RNA1_A, RNA1_C and RNA1_E were reacted with compound **11**.

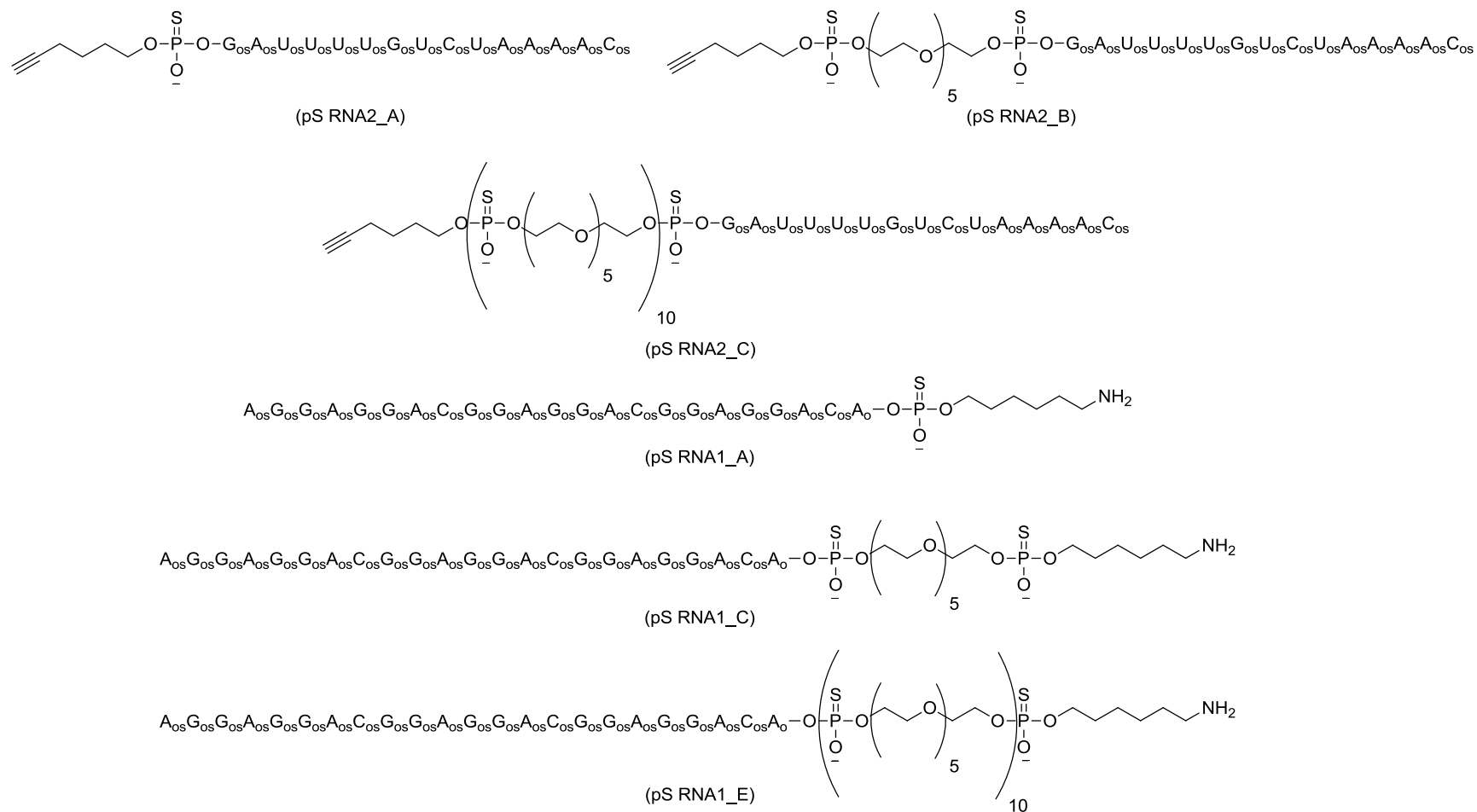


Figure 2.23. pS 2'OMe RNA sequences synthesised on an ABI 394 synthesiser using an “aged” PADS solution as an oxidiser.

2.4.9 – pS Trial Click Chemistry

pS RNA1_B, pS RNA1_D and pS RNA1_F were conjugated with the RNA2 series. Trial bioconjugation reactions (Table 2.9) producing tripartite sequences pS RNA3-pS RNA11 (Figure 2.24) were run to determine the attachment of the azide group and to investigate whether the pS 2'OMe RNA sequences behaved the same as the pS DNA sequences during click chemistry reactions.

Table 2.9. *Reactants and products for all pS 2'OMe RNA click chemistry bioconjugation reactions undertaken*

pS azide strands	pS alkyne strands	Product formed from bioconjugation
pS RNA1_B	pS RNA2_A	pS RNA3
pS RNA1_B	pS RNA2_B	pS RNA4
pS RNA1_B	pS RNA2_C	pS RNA5
pS RNA1_D	pS RNA2_A	pS RNA6
pS RNA1_D	pS RNA2_B	pS RNA7
pS RNA1_D	pS RNA2_C	pS RNA8
pS RNA1_F	pS RNA2_A	pS RNA9
pS RNA1_F	pS RNA2_B	pS RNA10
pS RNA1_F	pS RNA2_C	pS RNA11

The bioconjugated compound pS RNA3-pS RNA11 reactions were carried out as described in *Section 2.4.5* using 1eq of both the alkyne and azide and 100eq of CuTBTA. The reactions were desalted and analysed by 20% denaturing PAGE (Figure 2.25).

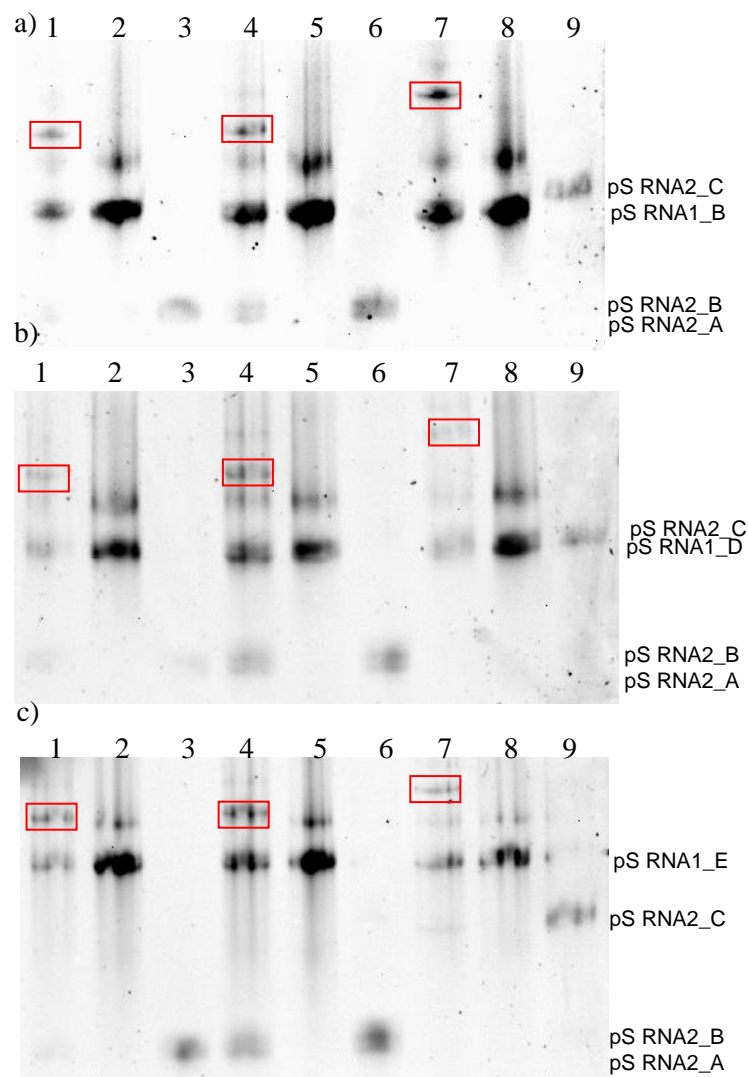


Figure 2.25. 20% denaturing polyacrylamide gels of click chemistry reaction samples. **a)** Click chemistry of pS RNA1_B with pS RNA2_A, pS RNA2_B and pS RNA2_C. Lane 1) Reaction between pS RNA1_B and pS RNA2_A. Lanes 2) pS RNA1_B. Lane 3) pS RNA2_A. Lane 4) Reaction between pS RNA1_B and pS RNA2_B. Lane 5) pS RNA1_B. Lane 6) pS RNA2_B. Lane 7) Reaction between pS RNA1_B and pS RNA2_C. Lane 8) pS RNA1_B. Lane 9) pS RNA2_C. **b)** Click chemistry of pS RNA1_D with pS RNA2_A, pS RNA2_B and pS RNA2_C. Lane 1) Reaction between pS RNA1_D and pS RNA2_A. Lane 2) pS RNA1_D. Lane 3) pS RNA2_A. Lane 4) Reaction between pS RNA1_D and pS RNA2_B. Lane 5) pS RNA1_D. Lane 6) pS RNA2_B. Lane 7) Reaction between pS RNA1_D and pS RNA2_C. Lane 8) pS RNA1_D. Lane 9) pS RNA2_C. **c)** Click chemistry of pS RNA1_F with pS RNA2_A, pS RNA2_B and pS RNA2_C. Lane 1) Reaction between pS RNA1_F and pS RNA2_A. Lane 2) pS RNA1_F. Lane 3) pS RNA2_A. Lane 4) Reaction between pS RNA1_F and pS RNA2_B. Lane 5) pS RNA1_F. Lane 6) pS RNA2_B. Lane 7) Reaction between pS RNA1_F and pS RNA2_C. Lane 8) pS RNA1_F. Lane 9) pS RNA2_C. Red boxes indicate product formed.

The red boxes highlighting product formation on Figure 2.25 show that the click chemistry reaction can occur on pS 2'OMe RNA strands with product conversions of ~30%. The product formation is harder to determine on these gels as the pS RNA1 series sequences show two bands for the starting materials. These two bands both represent the same material with the higher band representing the secondary structure of the specific RNA sequence (Figure 2.9 and 2.10).

2.4.10 – pS RNA6 & pS RNA7 Click Chemistry Scale Up

Scale up reactions were carried out as in *Section 2.4.6* using 1eq of alkyne and azide and 100eq of CuTBTA. The purification conditions were changed slightly to account for the higher T_m of the secondary structure (*Section 2.4.14*). To enable the purification of these strands a higher denaturing polyacrylamide gel percentage of 20% and longer gel running times were used. pS RNA6 was successfully purified on a 20% denaturing gel over 3 hours, however pS RNA7 samples were heated to 80°C prior to successful purified after 4¹/₂ hours (Figure 2.26)

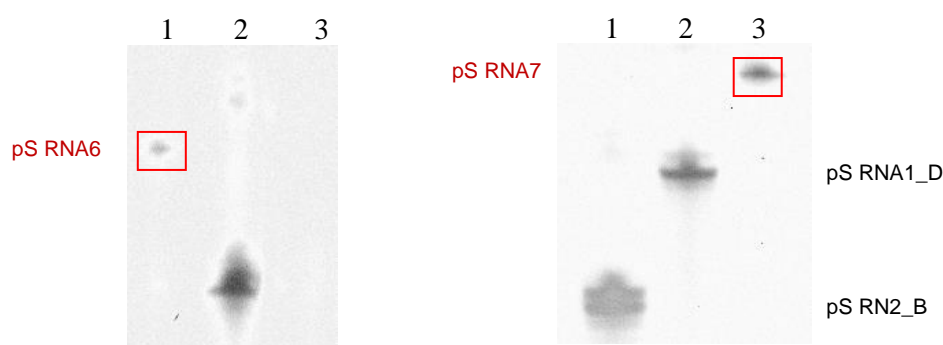


Figure 2.26. 20% denaturing polyacrylamide gel of pS RNA6 (a) and pS RNA7 (b) after purification. Gel A) Lanes 1-3) pS RNA6, pS RNA1_D and pS RNA2_A respectively. Gel B) Samples were heated to 80°C prior to loading. Lanes 1-3) pS RNA2_B, pS RNA1_D, pS RNA7 respectively. Products are in the red box.

2.4.11 – Solid Phase Synthesis of GGA Chemistry Strands

During TOES development a sequence called TOES 1 (GGA) was developed which incorporated pS at both ends with some pS in the middle of the strand (Figure 2.27). Addition of GGA to *SMN2* splicing reactions gave rise to a large increase in exon 7 inclusion levels when compared to GGA-O, but due to the potential of phosphorothioate possibly chelating to the Au nanoparticle (GNP) surface it was decided that GGA chemistry would not be used². However, as RNA6 and RNA14 provide a greater increase in exon 7 inclusion than GGA-O, GGA chemistry was investigated for use in tripartite sequences as these sequences would not be used on GNP. It was thought that using phosphorothioates within the backbone in RNA6 and RNA14 could potentially increase exon 7 inclusion levels to a higher level than GGA could. A GGA chemistry version of RNA14 was synthesised and named RNA16. A GGA chemistry version of RNA 6 could not be produced due to time constraints, but will be referred to as RNA6A. See Figure 2.27 for sequences of these molecules.

RNA16 was prepared using the same sequence as RNA14 but using RNA bases for the enhancer instead of 2'OMe RNA and having pS additions to cap the RNA to stop degradation. Due to time restraints the RNA6A version was not prepared. It would have given a direct comparison to RNA6, however work on this is on-going.

RNA16 was prepared by RNA synthesis on an ABI 394 synthesiser in 6 steps with PADS being used as the oxidiser in steps 1, 3 and 6 and iodine being used as the oxidiser in steps 2, 4 and 5 (*Section 2.5.10*). 2'OMe RNA and RNA bases were switched between step 4 and 5 (*Section 2.5.10*).

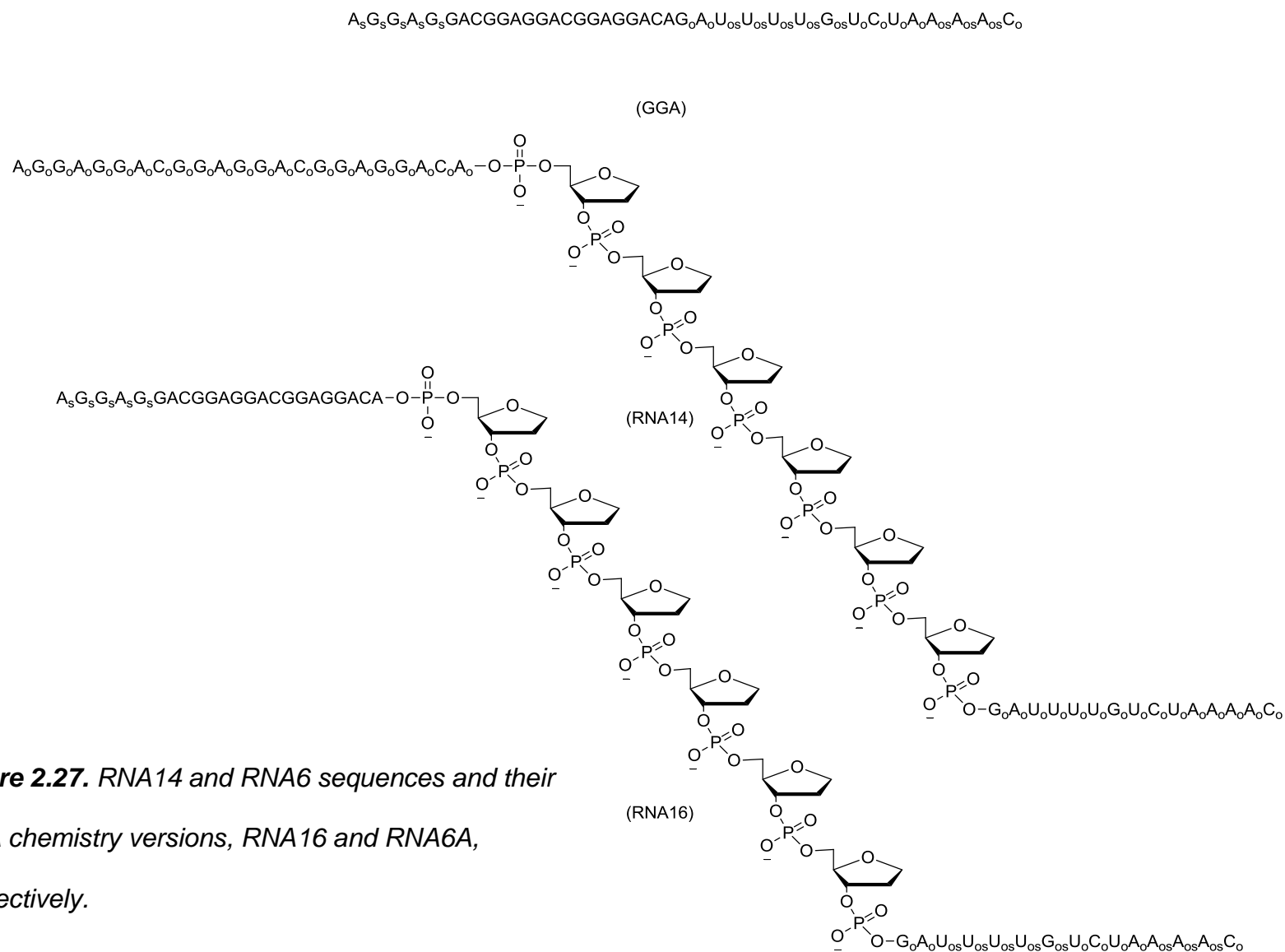


Figure 2.27. RNA14 and RNA6 sequences and their GGA chemistry versions, RNA16 and RNA6A, respectively.

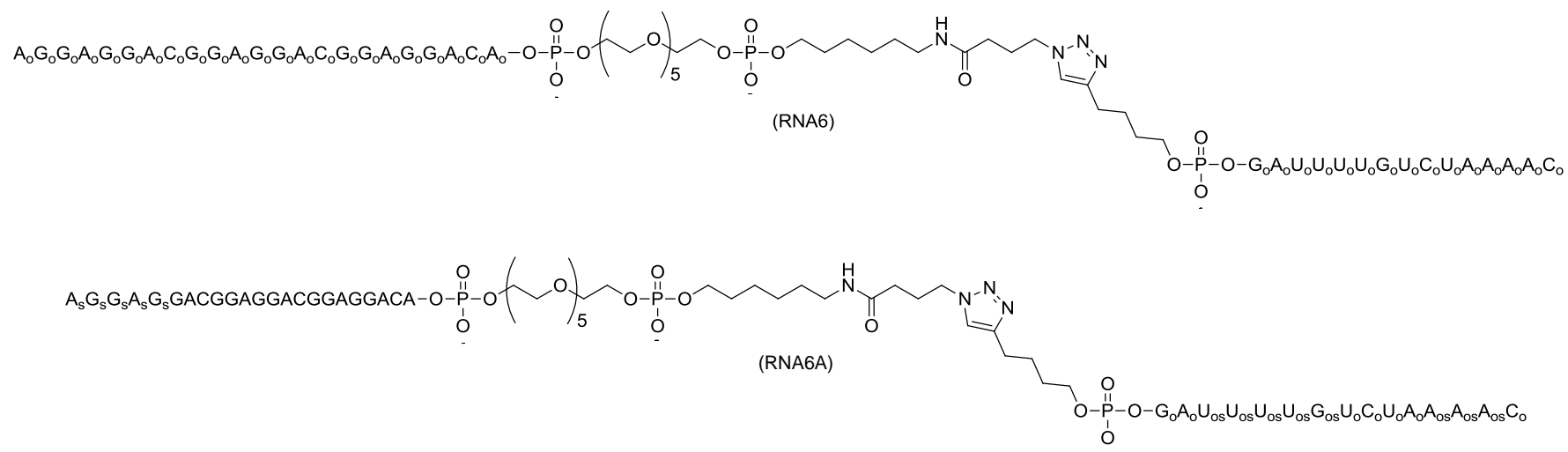


Figure 2.27. RNA14 and RNA6 sequences and their GGA chemistry versions, RNA16 and RNA6A, respectively (continued).

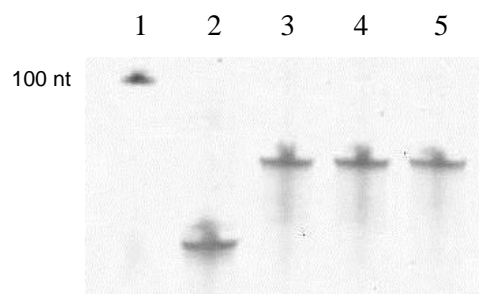


Figure 2.28. 20% denaturing polyacrylamide gel of RNA1_G, RNA14, RNA15 and RNA16 samples. Lane 1) DNA ladder. Lanes 2-5) purified RNA1_G, RNA14, RNA15 and RNA16, respectively.

As all tripartite sequences examined up to this point had formed secondary structures, it was assumed RNA16 would also form a secondary structure and so it was purified by RNA shadowing. (RNA shadowing is a technique that requires separation of samples by PAGE and uses short wavelength UV light to detect the RNA strand, which is then excised). The purified RNA16 was worked up using a standard elution protocol (*Section 2.5.1.5*) and analysed by 20% denaturing PAGE (Figure 2.28).

2.4.12 – Quadruplex Study

The UV-visible isothermal melting data shown in *Section 2.4.3.2* confirmed the presence of secondary structure within the RNA1 (enhancer sequence) with a T_m of ~54°C. As the enhancer sequence contains 3 cytosines, of which only two are predicted to be able to hydrogen bond (mfold, University at Albany) with guanines, a melting temperature of 54°C was thought to be too high for a standard hairpin conformation (Figure 2.29). The enhancer sequence does however contain large numbers of guanines, with approximately half the strand consisting of G bases, therefore guanine quadruplexes were investigated.

Sequences containing GGAGG repeats are known to form quadruplexes in both DNA and RNA. As the RNA1 enhancer sequence contains three repeats of this motif, further study was carried out to confirm the identity of the secondary structure observed. Investigations were carried out by Alex Cousin and Prof Mark Searle as they had the expertise and equipment required for examining quadruplexes. The secondary structure was investigated using NMR and Circular Dichroism (CD) with enhancer sequences RNA1, RNA18, RNA19 and pS RNA12 (Figure 2.30).

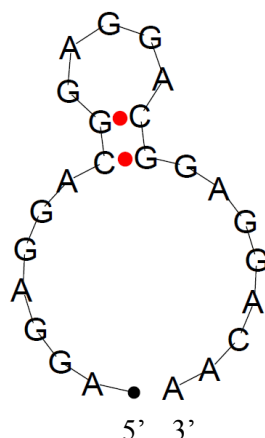


Figure 2.29. Secondary structure prediction for RNA1 sequence from a mfold program based at the RNA institute, College of Art and Science, University at Albany.

A_oG_oG_oA_oG_oG_oA_oC_oG_oG_oA_oG_oG_oA_oC_oG_oG_oA_oG_oG_oA_oC_oA_o

RNA1

AGGAGGACGGAGGACGGAGGACA

RNA18

AGGAGGACGGAGGAC

RNA19

A_{os}G_{os}G_{os}A_{os}G_{os}G_{os}A_{os}C_{os}G_{os}G_{os}A_{os}G_{os}G_{os}A_{os}C_{os}G_{os}G_{os}A_{os}G_{os}G_{os}A_{os}C_{os}A_o

pS RNA12

Figure 2.30. Strands used to investigate quadruplex secondary structures written 5' to 3' with _o and _s representing 2'OMe RNA and pS RNA, respectively.

Quadruplexes can adopt either a parallel or an anti-parallel conformation (Figure 2.52). Both conformations have been identified in DNA, but only the parallel conformation has been detected in RNA¹⁵².

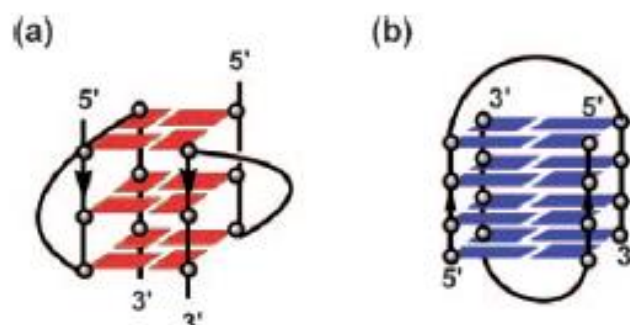


Figure 2.31. Intermolecular quadruplex structures found in nucleic acids. a) Parallel conformation. b) Anti-parallel conformation [diagram taken from Yan. *XJ*¹⁵³].

CD experiments can determine what type of quadruplex conformation exists by observing the light absorption of the strand. Parallel conformations have distinct absorption maxima at ~260nm, whereas anti-parallel strands have an absorption maxima at ~290nm, allowing for easy identification of conformation type¹⁵⁴. RNA1, RNA18, RNA19 and pS RNA12 were lyophilised before being re-dissolved in KCl 100mM, potassium phosphate (KP) 10mM, pH 7 in water to give final concentrations of ~ 4μM for the CD experiments. CD data was obtained after 0.2°C/minute cooling from 95°C to 25°C. All the strands gave characteristic absorption data for a parallel quadruplex (Figure 2.32).

Melting temperatures were also determined for all strands, with RNA strands RNA18 and RNA19 showing significant hysteresis between folding and melting experiments (Table 2.10).

Table 2.10. Melting temperatures sequences determined from CD spectra at 265nm running 0.2°C per minute. Fitted CD curves are shown in appendix (Figure 5.13)

Sequence Name	Sequence (5' → 3')	Strand Chemistry	T _m (°C) Folding	T _m (°C) Melting
RNA1	AGGAGGACGGAGGACGG AGGACA	2'OMe RNA	57.4	60.6
pS RNA12	AGGAGGACGGAGGACGG AGGACA	2'OMe/pS RNA	66.2	64.9
RNA18	AGGAGGACGGAGGACGG AGGACA	RNA	62.0	74.0
RNA19	AGGAGGACGGAGGAC	RNA	63.0	74.7

NMR experiments were carried out using the same buffer as the CD experiments, however, as NMR is less sensitive than CD, concentrations of 300µM-600µM were required. NMR experiments were undertaken for all the strands but the NMR spectra for RNA1, RNA18 and pS RNA12 were complicated due to the number of bases present in the strand (Figure 2.33). RNA19 however was a truncated version of the enhancer which was designed to reduce the complexity of the NMR spectra to aid assignment.

NMR resonance peaks indicative of a quadruplex are in the 10.5-11.6 ppm region which correspond to Hoogsteen base pairs between guanine imino and amino protons involved in a G quartet (Figure 2.34).

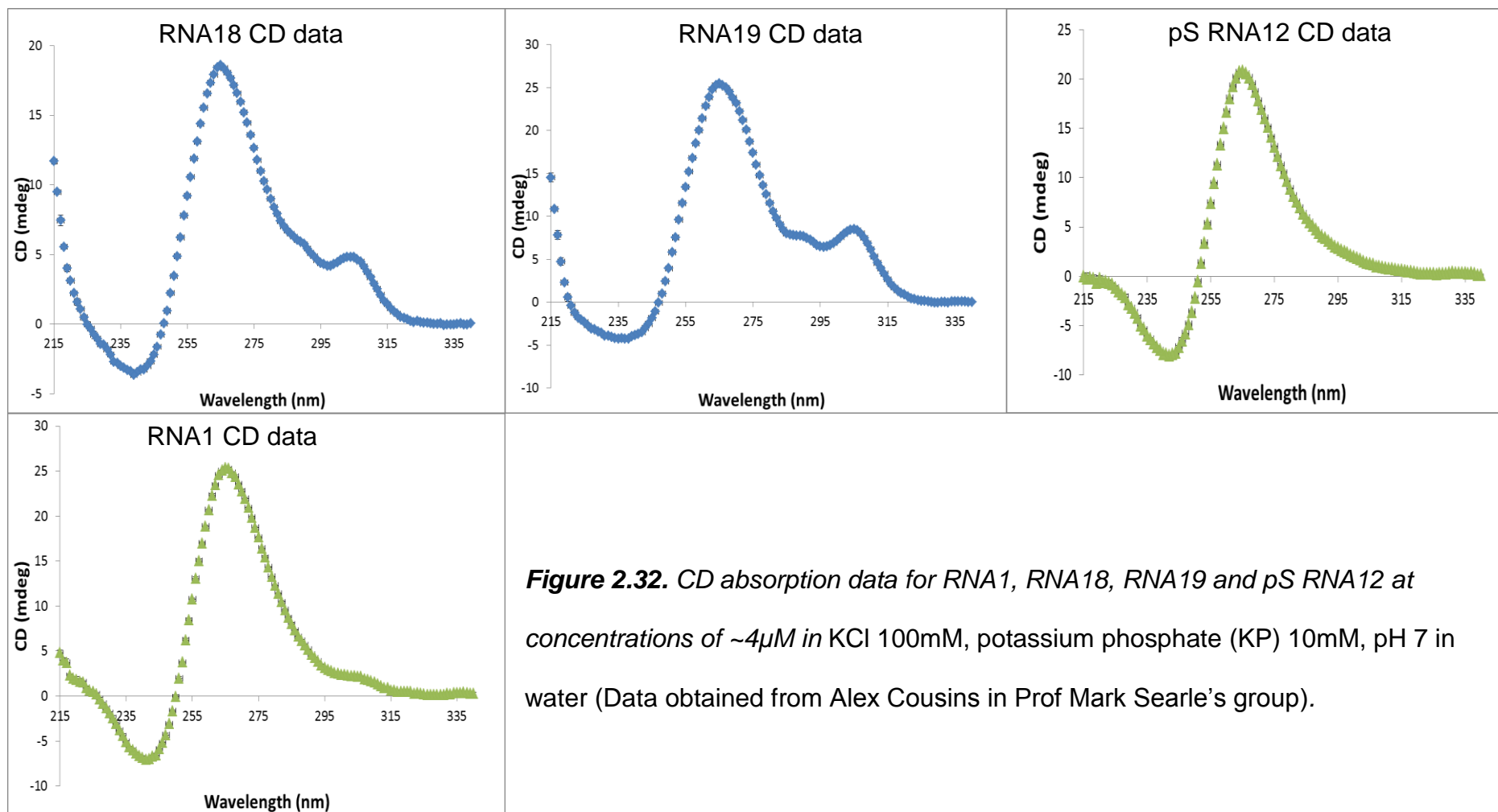


Figure 2.32. CD absorption data for RNA1, RNA18, RNA19 and pS RNA12 at concentrations of $\sim 4\mu\text{M}$ in KCl 100mM, potassium phosphate (KP) 10mM, pH 7 in water (Data obtained from Alex Cousins in Prof Mark Searle's group).

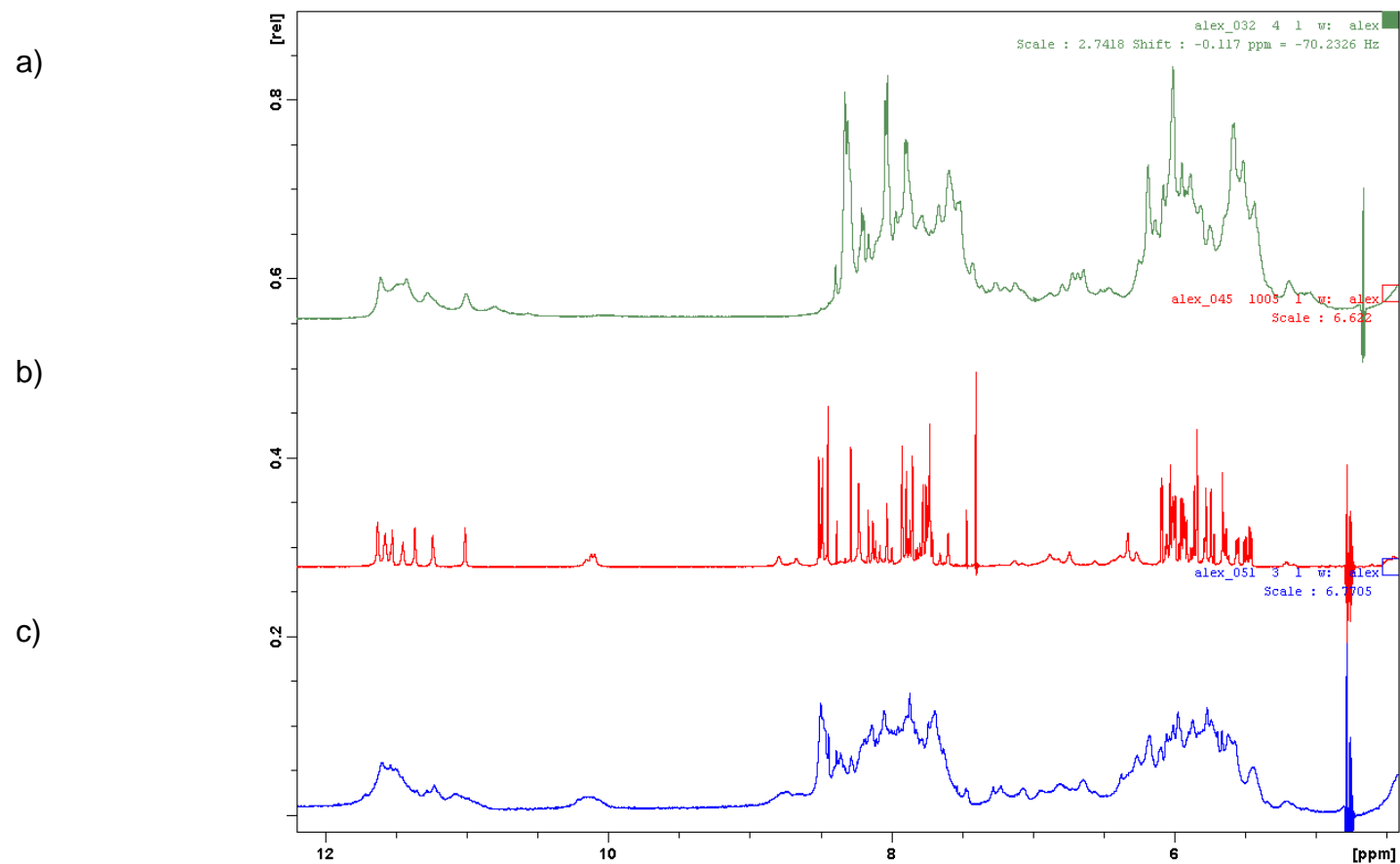


Figure 2.33.. NMR spectras. a) RNA1. b) RNA19. c) RNA18. NMR experiments were carried out at 25°C in KCl 100mM, potassium phosphate (KP) 10mM, pH 7 in water. NMR of pS RNA12 is included in the appendix (Figure 5.14) ((Data obtained from Alex Cousins in Prof Mark Searle's group).

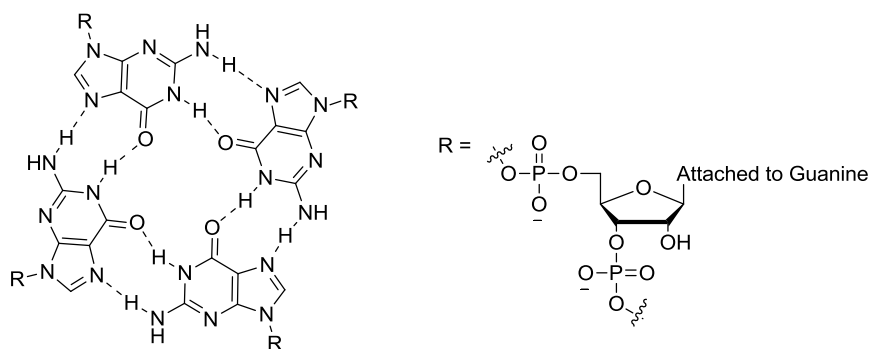


Figure 2.34. Hoogsteen base pairs in a G quartet [diagram adapted from Mergny, J.L]¹⁵⁵.

All the RNA strands examined showed the shifted imino protons, further confirming the presence of a G quadruplex structure within the enhancer sequence. Further experiments will be required to fully assign the quadruplex structure however RNA GGAGG repeats sequences have been shown to form a tetrad/hexad structure (Figure 2.35)^{154,156}.

The GGAGG repeat sequences has a small positive absorption at ~305nm which may be an indication of a hexad type quadruplex as this absorption is not observed in other G quadruplex structures¹⁵⁴. RNA1, RNA18 and RNA19 also contain an absorption at ~305nm suggesting that the tetra/hexad quadruplex structure could potentially be the quadruplex structure in the enhancer sequence.

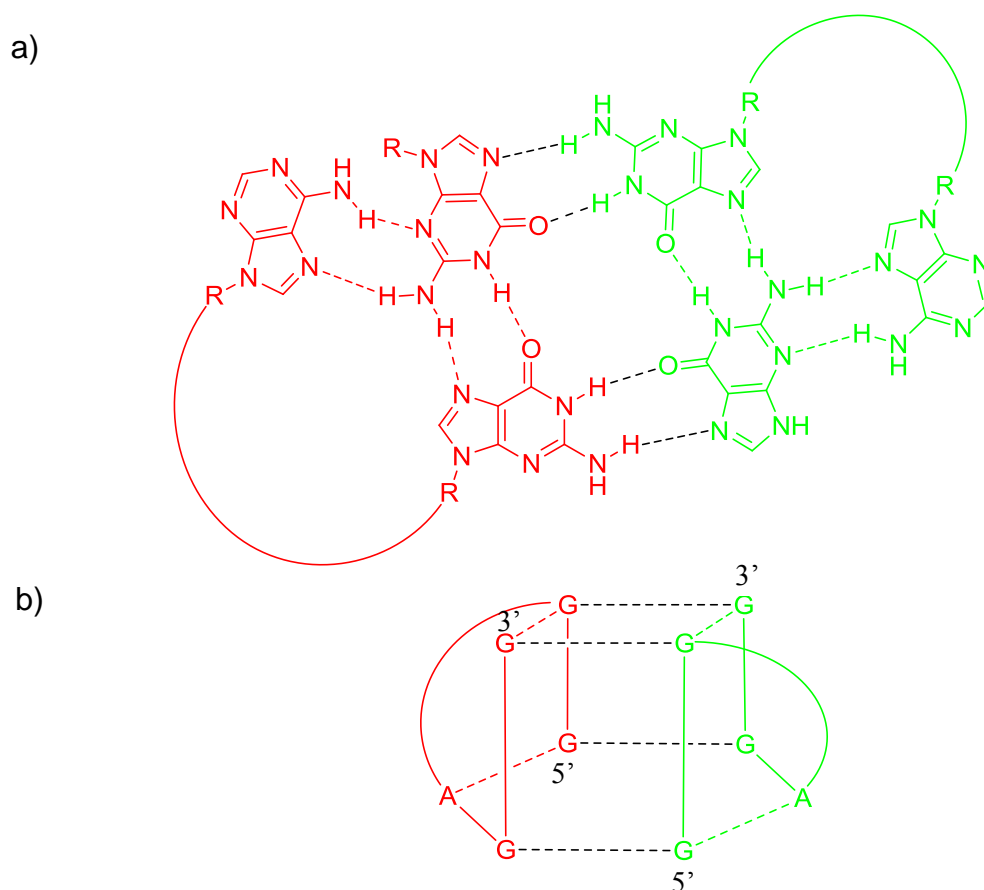


Figure 2.35. Tetrad/hexad structure formed by intermolecular GGA repeats. a) Hydrogen and bases involved in formation of the hexad. b) Schematic representation of the tetrad/hexad conformation. All dotted lines indicate hydrogen bonding, black dotted lines indicate intermolecular hydrogen bonding, red and green represent the individual strands [diagram adapted from Lipay, J.M]¹⁵⁴.

2.4.13 – Discussion

Preparation of compounds **10** and **11** enabled the attachment of the bioconjugation groups required for Cu mediated click chemistry. Compound **10** was incorporated during solid phase synthesis whereas compound **11** required post synthesis modifications of an amino modified strand.

During DNA/RNA synthesis extended coupling times were used for the RNA1 series with HEG units as it was thought that the flexibility of the HEG units would affect future coupling with RNA bases, therefore affecting strand purity. Purification of the 2'OMe RNA2 series was straightforward as RP-HPLC was used, but the 2'OMe RNA1 strands proved problematic as the purification by RP-HPLC could not provide separation (Figure 2.6). Ion exchange purification however was efficient for 2'OMe RNA1 strands with RNA1_C and RNA1_E sequences being successfully purified using this method (Figure 2.5). The ion exchanges gave two main peaks; the first being linear RNA1_C or RNA1_E, and the second caused by secondary structure formation in the sequences. (A T_m experiment on RNA1_A showed that secondary structure formation occurs in all sequences containing RNA1 (Figure 2.9)). The secondary structure formation in RNA1_C and RNA1_E was detected by PAGE (Figure 2.7), but successfully denatured by heating at 80°C (Figure 2.10). Experiments using CD and NMR (by Alex Cousins and Prof Mark Searle) revealed the secondary structure that forms in sequences containing RNA1 is a parallel quadruplex (Figure 2.32). The comparatively large differences seen between the melting temperatures calculated from CD folding and melting spectra for RNA strands (Table 2.10), RNA18 and RNA19, suggest that the quadruplex structure may be intermolecular rather than unimolecular. This is because intermolecular interactions take longer to form than unimolecular interactions, therefore if the quadruplex structure was unimolecular, the CD folding and melting spectra would most likely have shown much closer melting temperatures (Table 2.10). Further experiments are required to determine the conformation of the G

quadruplex however previous experiments with the same GGAGG motif suggest the presence of a tetra/hexad quadruplex structure (Figure 2.35).

As click chemistry on 2'OMe RNA strands had not been tried before two different concentrations of Cu were trialled; 10eq and 100eq (Figure 2.14). Only the reactions with 100eq of CuTBTA formed product, therefore 100eq of CuTBTA was used for the subsequent reactions. An interesting observation, in both short and long click chemistry reactions, was the apparent lack of further reaction after 15 minutes, even in the presence of more starting material. To investigate why this occurred several different reaction conditions were trialled and are summarised in Table 2.11.

From the Table 2.11 it can be seen that increasing alkyne concentrations, Cu concentrations or the addition of ascorbate seemed to have no effect on product conversion. This shows that the reaction had stopped due to Cu oxidation or alkyne levels as originally proposed. In the two reactions where ascorbate was present, the RNA1 strand seemed to disappear. However no increase in product levels was detected for either reaction and it was therefore proposed that high ascorbate concentrations in the presence of Cu causes RNA1 degradation. Since none of the conditions trialled increased product conversion the click chemistry reactions were carried out in the same way as the original prep, using 1eq of both alkyne and azide and 100eq of CuTBTA. To increase product yields from the click chemistry reactions, templates could have been used, however due to the large number of HEG units incorporated in strands RNA1_F and RNA2_C, a long template strand of ~80nt would have been required¹⁴⁹. This would have been very costly and problematic to synthesise therefore template directed click chemistry was not considered. However once

the optimal linker and spacing distance between the strands had been determined in the *SMN2* splicing assay then templates could be made to increase the yields for the bioconjugation reactions for developments in rectifying other splicing defects that causes disease pathways.

Table 2.11. *Conditions trialled to improve product conversion during Cu-catalysed Huisgen [3+2] click chemistry*

Conditions trialled	Outcome
Increasing alkyne concentration	No increase in product conversion
Several additions of Cu	No increase in product conversion
Addition of ascorbate to reactions	No increase in product conversion – Reactions with 300eq of CuTBTA and 600eq of ascorbate showed RNA1 degradation.
Adding Denaturing agents to the reaction to melt secondary structure	No increase in product conversion
Increasing concentration by evaporation	No increase in product conversion - RNA1 strand used up but as no increase in product levels assume RNA1 strand just degraded.

When investigating how to purify strands RNA3-11 and RNA15, a trial reaction was run in which a standard scale up click chemistry reaction was split into two samples. One sample was purified by ion exchange and the other by PAGE. PAGE purification achieved a significantly higher yield of 3% compared to just 0.3% from ion exchange. Purification of strands RNA3-11 and RNA15 were then carried out by gel electrophoresis giving yields of ~10%. These yields are

very low, but when considering that the reaction never went to completion, that only small amounts of products were required for *in vitro* tests and that the unreacted starting materials could be re-used in further click chemistry reactions, the low yields did not pose a problem.

Phosphothioate backbones (pS) had been shown to increase protein binding² which could be beneficial for the tripartite sequences, however phosphothioate nucleic acid Cu-catalysed Huisgen [3+2] click chemistry reactions had not been carried out before. To investigate click chemistry compatibility with phosphothioates, DNA versions of TOES were synthesised. The phosphothioate versions were prepared using PADS oxidiser which after “ageing” for 2 days provided 99.9% sulphur addition¹⁵⁷. The benefit of using PADS over other phosphothioate oxidisers is that it is inexpensive but still maintains high levels of sulphur incorporation required for efficient synthesis.

pS click chemistry reactions behaved exactly the same as the non-pS click chemistry reactions and also stalled by the 15 minute time point. The secondary structures formed by the pS RNA1 series were stronger than those formed by the 2'OMe RNA1 sequences, having T_m of 66-64°C and 60-57°C, respectively (Table 2.9). Due to the high T_m of the pS RNA1 series, higher percentage gels and heat treatment was required to obtain pure samples (Figures 2.26). pS version of RNA6 and RNA7 were prepared after initial splicing results showed that these two strands benefitted exon 7 inclusion levels (*Section 3.5.1.2*). If the activity of these strands could be increased by adding a pS backbone it would be an easy way to optimise the system.

During TOES development a sequence called TOES 1 (GGA) was developed which showed an increase in exon 7 inclusion levels from 1.41 to 1.90 when compared to GGA O (TOES 27). As both RNA6 and RNA14 (*Section 3.5.1.6*) caused an increase in exon 7 inclusion compared to GGA-O, addition of the spacers into GGA could potentially boost exon 7 inclusion levels even higher. Unfortunately due to time constraints only RNA16 was made which had the same sequence as RNA14 except that the enhancer was RNA not 2'OMe RNA and phosphothioates were added to the strand.

2.4.14 – Conclusion

Using Cu catalysed click chemistry, a new conjugation method was developed for conjugating 2'OMe RNA and 2'OMe RNA pS strands together. Coupling yields for these reactions were low due to the reactions not going to completion. Once the optimal tripartite sequence has been determined for *SMN2* splicing, these yields could be increased by making template strands.

Enhancer sequence (RNA1) has been found to have a parallel quadruplex secondary structure. This structure would appear to be intermolecular and may be a tetrad/hexad type structure, as other sequences similar to RNA1 have this type of structure. Further studies are required to determine the exact structure of the parallel quadruplex.

2.5 – Experimental

UV-Vis measurements were taken with a Thermo-Scientific Nanodrop 1000.

2.5.1 – General Procedures

2.5.1.1 – Standard DNA/RNA Synthesis

Solid phase DNA/RNA synthesis was carried out on a DNA/RNA ABI 394 synthesiser using standard reagents and phosphoramidites for synthesis purchased from Link technologies. 5- Benzythio-1-H-tetrazole in anhydrous acetonitrile (0.3M) was used as the activator as provides superior coupling yields over tetrazole¹⁵⁸. THF/pyridine/acetic anhydride (8:1:1) Cap A and 10% Methylimidazole in THF Cap B were used to acetylate unreacted 5' hydroxyl groups arising from incomplete phosphoramidite couplings. Iodine in THF/pyridine/water (7:2:1, 0.02 M) was used as the oxidiser and 3% Trichloroacetic acid in DCM used to deprotect the DMT group of the 5' hydroxyl.

2.5.1.2 – HPLC Protocol

Analytical HPLC was run on a Phenomenex® Clarity Oligo-RP™ column (250 x 4.6mm) running a linear gradient from 5-50% 0.1M TEAA in water/0.1M TEAA 80:20 ACN:water at 1ml/minute over 20 minutes. An examples gradient for HPLC separation is shown in Figure 2.36.

Semi prep HPLC was run on a Phenomenex® Clarity Oligo-RP™ column (250 x 10mm) running a linear gradient from 5-60% 0.1M TEAA in water/0.1M TEAA 80:20 ACN:water at 4ml/minute over 25 minutes.

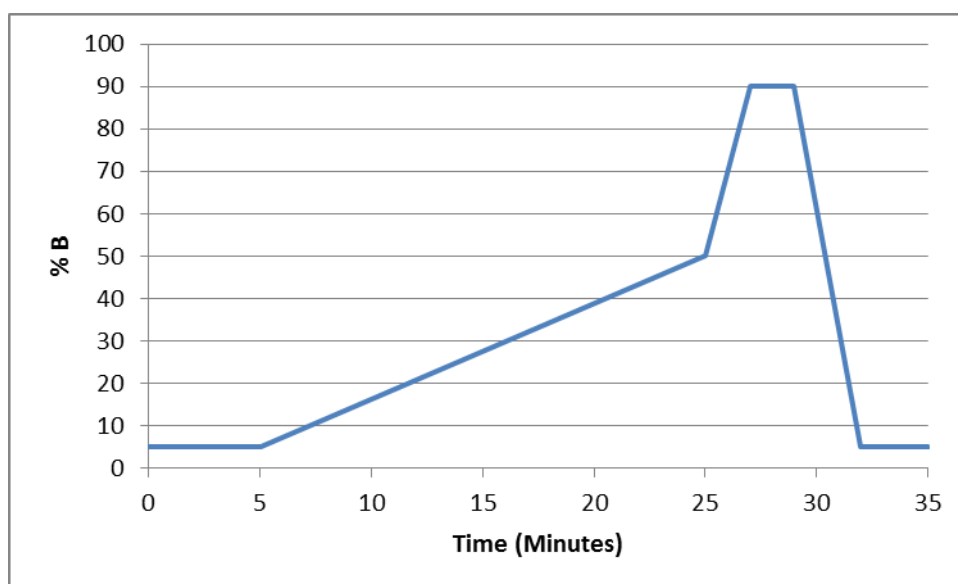


Figure 2.36. Gradient used for HPLC analysis of DNA and 2'OMe RNA sequences from 5-50% 0.1M TEAA in water/0.1M TEAA 80:20 ACN:water at 1 ml/minute over 20 minutes.

2.5.1.3 – Gel Imaging

Analysis of DNA and RNA prepared by solid phase synthesis were visualised by staining with either SYBR[™] Gold or SYBR[™] Green and imaged using a VersaDoc[™] MP imaging system.

2.5.1.4 – Desalting

Small scale desalting was carried out using MF-Millipore[™] membranes. These membranes consist of mixed cellulose ester which are hydrophilic and contain 0.025µm pores. Large scale desalting was carried out using GE Healthcare Illustra NAP[™] columns which were packed with DNA grade G 25 Sephadex[™].

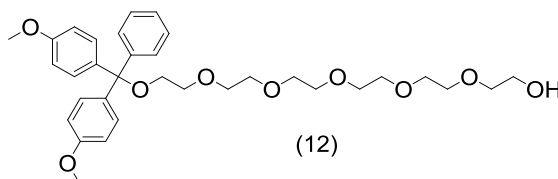
2.5.1.5 – Click Chemistry Scale Up Protocol

A stock vortexed solution of RNA-azide (60µl, 17.2µM) and an RNA-alkyne (60µl, 17.2µM) was aliquoted into 12 tubes (10µl per tube). CuBr (3.1mg, 2.16×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 216µl,

0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. CuTBTA (1µl, 8.62mM) was then added to each aliquot, vortexed and then heated to 40°C at 750rpm for 15mins. The samples were desalted (MF-Millipore™ Membrane, 20mins) and purified by gel electrophoresis using 10% denaturing polyacrylamide gels at 100V for 1hour (four gels used in total). The DNA/RNA within the reaction mixtures were identified using SYBR™ gold, image captured on a VersaDoc™ imager and the relevant bands isolated by excision from the gel. Each excised band was placed in separate tubes, treated with RNA elution buffer (350µl per tube) and left at 5°C overnight. The samples were then heated to 30°C for 10 minutes to re-dissolve any precipitated SDS before transferring the supernatants to new tubes. Ethanol (1.05ml, 3 x vol of RNA elution buffer) was then added to the supernatants, vortexed and finally centrifuged for 30 minutes at 13,400rpm. The supernatants were removed and the resultant RNA pellet washed with ethanol (250µl) before being centrifuging again for 15 minutes at 13,400rpm. The supernatants were once again removed and the pellets dried, re-suspended in water and combined to afford the final click products.

2.5.2 – Synthesis of 2'OMe RNA Conjugates

2.5.2.1 – Preparation of 1,1-bis(4-methoxyphenyl)-1-phenyl-2,5,8,11,14,17-hexaoxonadecan-19-ol (12)

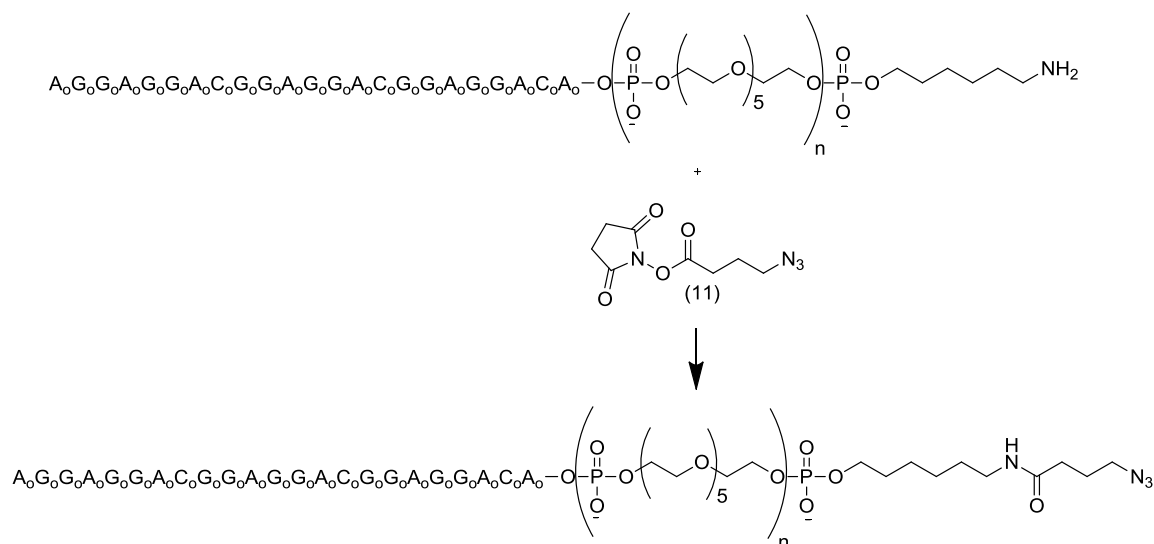


Compound **12** was prepared according to the protocol of M. Ma *et al*⁵. A solution of hexa(ethylene glycol) (1.78ml, 7.08mmol), DMAP (57.7mg, 4.72x10⁻⁴mol) in anhydrous pyridine (35ml) was transferred into a round bottom flask with molecular sieves (4Å) and left stirring under nitrogen for one hour. A solution of DMT-Cl (1.6g, 4.72mmol) in pyridine (5ml) was then added to the reaction mixture drop-wise over 30 minutes and then left stirring for 3 days. Methanol (10ml) was then added to the reaction mixture stirred at RT for 10 minutes and the reaction mixture concentrated *in vacuo*. Chloroform (30ml) was then added to the reaction mixture, washed with 5% NaHCO₃ (30ml) and brine (30ml) twice. The aqueous layers were then combined and re-extracted with chloroform (100ml). The organic layers were combined and dried with magnesium sulphate and the solid filtered. The resultant solution was then concentrated *in vacuo*. Column chromatography (SiO₂, 1:1 ethyl acetate/petroleum ether with 1% pyridine → 100% ethyl acetate afforded compound (**12**) as a yellow oil in 29 % (1.19 g); ES/MS: m/z 583, ¹H NMR (400 MHz; C₆D₆) δ: 3.30-3.60 (m, 30 H) 6.74-6.79 (m, 4 H) 7.07 (t, *J* = 7.32 Hz, 1 H) 7.19 (t, *J* = 7.77 Hz, 2 H) 7.53 (d, *J* = 8.87 Hz, 4 H) 7.70 (d, *J* = 8.44 Hz, 2 H).

2.5.2.2 – Preparation of Bioconjugation Reactants for Click Chemistry

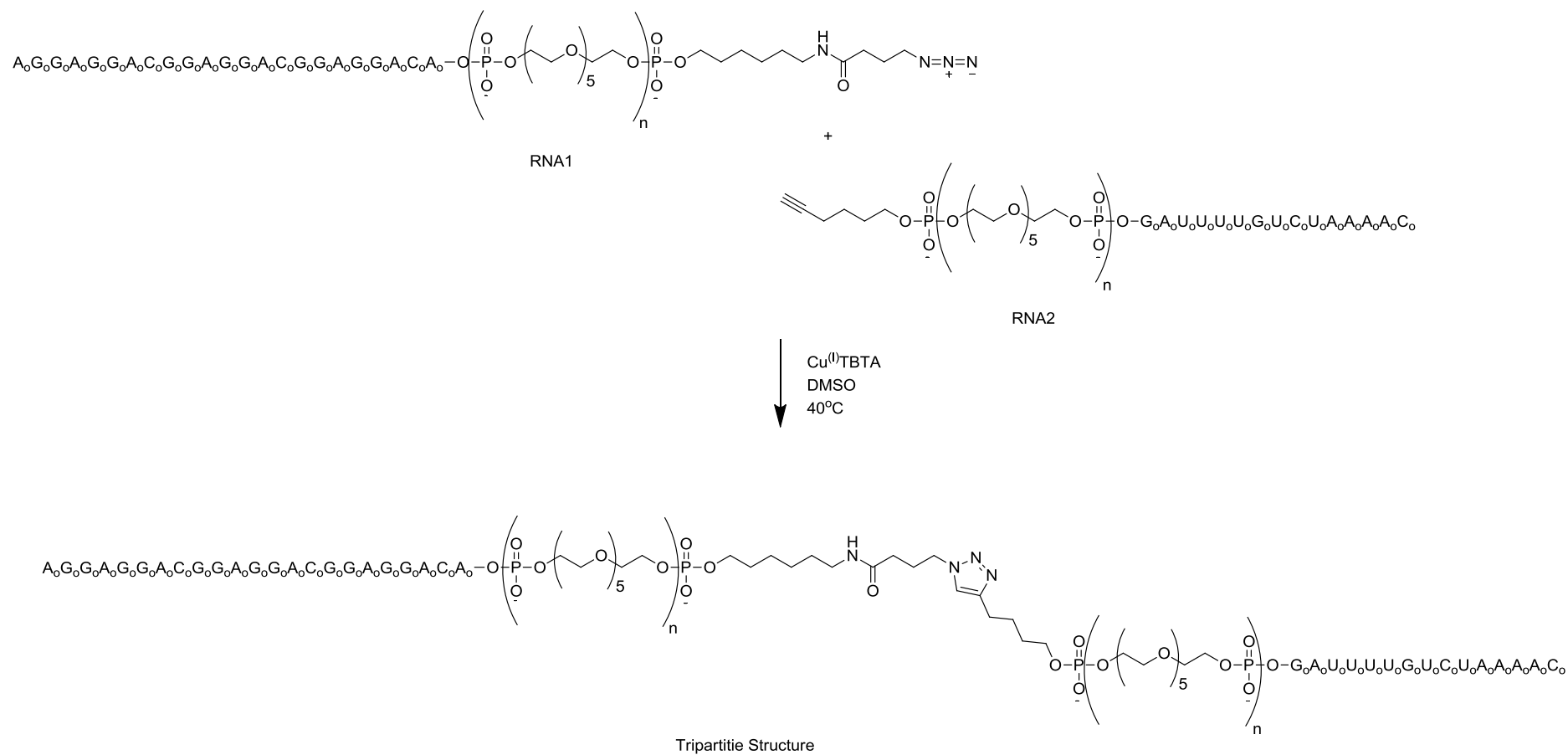
Scheme 2.7 below highlights the synthetic route into the tripartite structures. The RNA 1 series consist of an enhancer sequence with an azide and the RNA2 series consist of an annealing sequence with an alkyne for an *SMN2* mini gene. Different numbers of compound **13** units were bound between an RNA sequence (Table 2.12) and either an azide (RNA1 series) or an alkyne (RNA2

series). The RNA1 series was produced through solid phase synthesis with an amino modified linker which then underwent NHS coupling with azide (compound **11**) (Section 2.4.2) to produce azides RNA1_B, RNA1_D and RNA1_F (Scheme 2.6).



Scheme 2.6. Schematic of an NHS conjugation reaction between compound **11** and RNA 1 sequences RNA1_A, RNA1_C and RNA1_E used in bioconjugation reactions (Sections 2.5.2.11 - 2.5.2.13) _o represents 2'OMe RNA.

RNA 1 and RNA 2 series were conjugated together using Cu-catalysed click chemistry as shown in Scheme 2.7 to produce tripartite structures (Table 2.12).

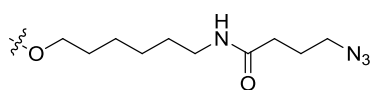


Scheme 2.7. Schematic of a Cu-catalysed Huisgen [3+2] cycloaddition reaction between the RNA 1 series and the RNA 2 series to produce tripartite structures RNA3-RNA11 used in splicing reactions (Section 2.5.3.9). \circ represents 2'OMe RNA.

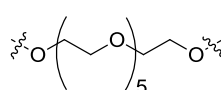
Table 2.12. RNA1_A, RNA1_C, RNA1_E, RNA2_A, RNA2_B and RNA2_C prepared using the ABI synthesizer (Sections 2.5.2.8 and 2.5.2.9). Compounds RNA 1_B, RNA 1_D and RNA 1_E formed from the NHS coupling reaction of RNA1_A, RNA1_C and RNA1_E with compound 4 (Sections 2.5.2.11 – 2.5.2.13). W, X, Y, and Z are shown below and _o represents 2'OMe RNA

Sequence Name	Sequence 5' to 3'
RNA2_A	YG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o
RNA2_B	YXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o
RNA2_C	YXXXXXXXXXXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o
RNA1_A	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o Z
RNA1_B	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o W
RNA1_C	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XZ
RNA1_D	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XW
RNA1_E	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XXXXXX XXXXXZ
RNA1_F	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XXXXXX XXXXXW

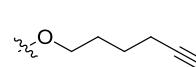
W =



X =



Y =



Z =

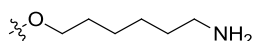
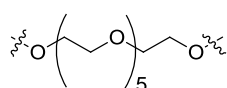


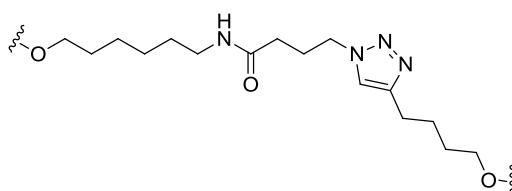
Table 2.13. Products formed from click chemistry of starting materials. X and Z structure are given below and _o represents 2'OMe RNA

Sequence Name	Product 5' – 3'	Starting sequences
RNA3	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o ZG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o	RNA 1_B RNA 2_A
RNA4	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o ZXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o	RNA 1_B RNA 2_B
RNA5	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o ZXXXXXXXXXXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o	RNA 1_B RNA 2_C
RNA6	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XZG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o	RNA 1_D RNA 2_A
RNA7	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XZXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o	RNA 1_D RNA 2_B
RNA8	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XZXXXXXXXXXXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o	RNA 1_D RNA 2_C
RNA9	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XXXXXXXXXXXXZG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o	RNA 1_F RNA 2_A
RNA10	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XXXXXXXXXXXXZXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o	RNA 1_F RNA 2_B
RNA11	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XXXXXXXXXXXXZXXXXXXXXXXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o	RNA 1_F RNA 2_C

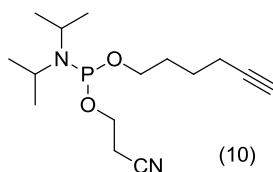
X =



Z =



2.5.2.3 – Preparation of 2-cyanoethyl N, N diisopropyl hex-5-yn-1-yl phosphoramidite (**10**)¹⁵⁹



A solution of hexyn-1-ol (225 μ l, 2.04 $\times 10^{-3}$ mol), diisopropylethylamine (462 μ l, 2.65 $\times 10^{-3}$ mol) in degassed anhydrous DCM (2ml) was added to a flame dried, argon filled flask. The flask was evacuated and re-filled with argon before the addition of a solution of chlorophosphoramidite (683 μ l, 3.06 $\times 10^{-3}$ mol) in dry DCM (2ml). The reaction mixture was stirred at RT for 2 hours. The product was purified by column chromatography using neutralised SiO₂ (neutralised with pyridine and heated in an oven overnight at 120°C; eluent 30% dry ethyl acetate/ dry hexane with 1% pyridine). Dried 100ml round bottom flasks were used to collect fractions, collecting five 50ml fractions. The solvent was removed by direct evaporation on a high vacuum line fitted with a cold trap. Each fraction was analysed by ¹H and ³¹P NMR. Compound **10** was afforded as a colourless oil in 83% (504mg); ¹H NMR (300MHz; C₆D₆) δ : 1.03 (t, J = 6.45 Hz, 12 H) 1.34-1.54 (m, 4 H) 1.57-1.69 (m, 3 H) 1.85-1.93 (m, 2 H) 3.11-3.28 (m, 2 H) 3.33-3.56 (m, 4 H), ³¹P NMR (161MHz, C₆D₆) δ : 147.96 (s, 1 P).

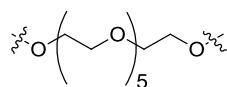
2.5.2.4 – Solid Phase Synthesis on a 0.2 μ mol for RNA1 and RNA2 Series

2'OMe RNAs were prepared by the DMT- and β -(cyanoethyl) phosphoramidite method on 0.2 μ mol CPG supports using an Applied Biosystems 394 machine. Coupling times were elongated to 11 minutes for all bases.

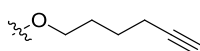
Table 2.14. Products RNA1_A, RNA1_C, RNA1_E, RNA2_A, RNA2_B and RNA2_C formed by solid phase synthesis using an ABI 394 synthesiser. X, Y and Z structures are shown below and _o represents 2'OMe RNA bases

Sequence Name	Sequence – 5' to 3'
RNA2_A	YG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o
RNA2_B	YXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o
RNA2_C	YXXXXXXXXXXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o
RNA1_A	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o Z
RNA1_C	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XZ
RNA1_E	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XXXXXX XXXXZ

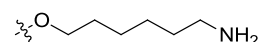
X =



Y =



Z =



Upon completion of the automated synthesis the sequences were cleaved from the solid support by aminolysis (35% aqueous ammonia) for 4 hours at RT. Deprotection of all protecting groups was achieved by incubation of the 35% aqueous ammonia solution at 55°C for 16 hours. The aqueous ammonia was removed by evaporation and the RNA2 series purified by reverse phase HPLC (Section 2.5.1.2).

Reverse phase (RP) HPLC purification of the RNA1 series was not possible due to the presence of a G-Quadruplex structure (Section 2.4.14). RNA1_C and RNA1_E were purified by ion exchange (10mM HEPES, 20mM NaCl to 10mM HEPES, 1M NaCl over 30 minutes for RNA1_C and 80 minutes for RNA1_E).

The purification success was determined by 10% denaturing PAGE (Section 2.4.3.2, Figure 2.8).

Table 2.15. Characterisation summary for the RNA products RNA1_A, RNA1_C, RNA1_E, RNA2_A, RNA2_B and RNA2_C prepared by solid phase synthesis. MALDI-TOF, HPLC and Ion exchange data in (Section 2.4.3.2, Figures 2.11- 2.14 and appendix (5.8 a-x)). Products were combined with the 1 μ mol synthesis before undergoing MALDI-TOF

Sequence	Progress	Amount (μ g)	Expect MW	MALDI-TOF MW	RP-HPLC analytical (RT)	Ion Exchange (RT)
RNA1_A	Abandoned	-	8131.29		-	-
RNA1_C	Purified	2.5	8475.58	8476.28	-	15.3 (16-20 for quadruplex)
RNA1_E	Purified	2.8	11574.15	11575.21	-	20.3 (26-33 for quadruplex)
RNA2_A	Purified	18.9	5092.23	5092.82	15.500	-
RNA2_B	Purified	5	5436.51	5436.76	16.148	-
RNA2_C	Purified	13.9	8535.08	8535.52	16.548	-

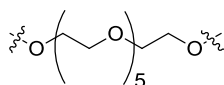
2.5.2.5 – Solid Phase Synthesis on a 1 μ mol for RNA1 and RNA2 Series

2'OMe derivatives were prepared by the DMT- and β -(cyanoethyl) phosphoramidite method on 1 μ mol CPG supports using an Applied Biosystems 394 machine. A double coupling protocol and extension of coupling times to 18 minutes was applied for strands RNA1_C and RNA1_E with coupling times for the other strands at 11 minutes.

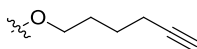
Table 2.16.. Products RNA1_A, RNA1_C, RNA1_E, RNA2_A, RNA2_B and RNA2_C formed by solid phase synthesis using an ABI 394 synthesiser. X, Y and Z structures are shown below and _o represents 2'OMe RNA bases

Sequence Name	Sequence – 5' to 3'
RNA2_A	YG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o
RNA2_B	YXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o
RNA2_C	YXXXXXXXXXXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o A _o C _o
RNA1_A	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o Z
RNA1_C	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XZ
RNA1_E	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XXXXXX XXXXZ

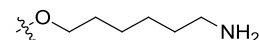
X =



Y =



Z =



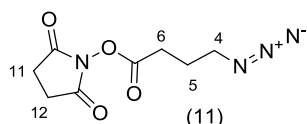
Upon completion of the automated synthesis the sequences were cleaved from the solid support by aminolysis (35% aqueous ammonia) for 4 hours at RT. Deprotection of all protecting groups was achieved by incubation of the 35% aqueous ammonia solution at 55°C for 16 hours. The aqueous ammonia was removed by evaporation and the RNA2 series purified by reverse phase HPLC (Section 2.5.1.2).

RNA1 series were desalted using Illustra NAP™ 25 columns, lyophilised and re-dissolved in ddH₂O to work up the strands. Products were analysed using 10% denaturing PAGE (Section 2.4.3.2, Figure 2.11) and confirmed by MALDI-TOF (Table 2.17).

Table 2.17. Characterisation summary for the RNA products RNA1_A, RNA1_C, RNA1_E, RNA2_A, RNA2_B and RNA2_C prepared by solid phase synthesis. Section 2.4.3.2, Figures 2.11- 2.14 and appendix (5.8). Products were combined with the 0.2 μ mol synthesis before undergoing MALDI-TOF

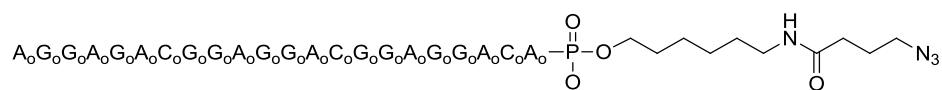
Sequence	Progress	Amount (mg)	Expect MW	MALDI-TOF MW	RP-HPLC analytical (RT)
RNA1_A	Desalted	0.61	8131.29	8131.83	-
RNA1_C	Desalted	1.36	8475.58	8476.28	-
RNA1_E	Desalted	1.74	11574.15	11575.21	-
RNA2_A	Purified	1.02	5092.23	5092.82	15.253
RNA2_B	Purified	1.28	5436.51	5436.76	16.292
RNA2_C	Purified	1.80	8535.08	8535.52	16.682

2.5.2.6 – Preparation of 2,5-dioxopyrrolidin-1-yl 4-azidobutanoate (**11**)¹⁴⁸



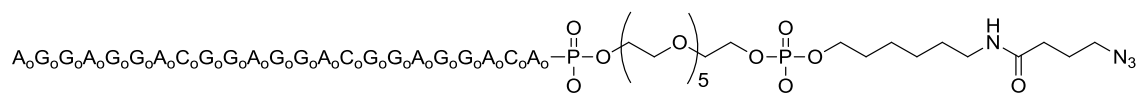
To a solution of 4-azidobutanoic acid (50mg, 3.87x10⁻⁴mol) in anhydrous DCM (2ml) solid N-hydroxysuccinimide (66.9mg, 5.81x10⁻⁴mol) and HOBt (78.5mg, 5.81x10⁻⁴mol) were added. The reaction mixture was cooled to 0°C under an argon atmosphere before EDC (111.4, 5.81x10⁻⁴mol) was then added and the reaction left to warm overnight. The product was purified by column chromatography (SiO₂, eluent 1:1 ethyl acetate/petroleum ether) to afford compound (**11**) as a white solid in 34 % (29.8mg); ¹H NMR (300MHz; CDCl₃) δ : 1.94 (q, J = 6.83 Hz, 2 H, C5) 2.66 (t, J = 7.19 Hz, 2 H, C6) 2.77 (s, 4 H, C11 & C12) 3.38 (t, J = 6.53 Hz, 2 H, C4).

2.5.2.7 – Preparation of RNA1_B from RNA1_A and Compound 11



To a solution of RNA1_A (14μg in 100μl water) in aqueous sodium hydrogen carbonate (30μl, 0.025M, pH~8) was added a solution of compound **11** (2mg, 8.84x10⁻⁶mol) in DNA-grade acetonitrile (70μl). The mixture was vortexed and a precipitate formed, as these samples had been desalted before this reaction, sodium hydrogen carbonate (40μl, 0.025M, pH~8) was added drop-wise until the precipitate re-dissolved before leaving the reaction overnight at RT. The product was concentrated to remove the acetonitrile and desalted through an Illustra NAP™ 25 column. The mass spectrum of RNA1_B gave the expected mass of [M+H] = 8243.63 (Appendix Figure 5.9).

2.5.2.8 – Preparation of RNA1_D from RNA1_C and Compound 11

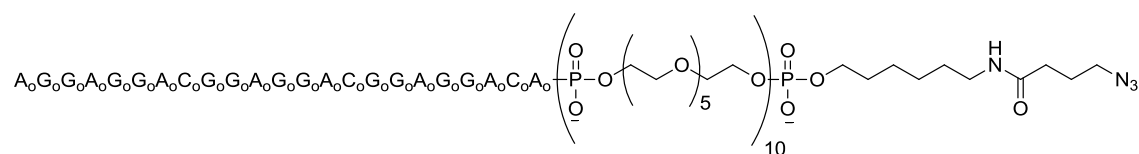


To a solution of RNA1_C (1.6μg in 66.7μl water) in aqueous NaHCO₃ (180μl, 0.025M, pH~8) was added a solution of compound **11** (1mg, 8.84x10⁻⁶mol in DNA-grade acetonitrile (420μl). The mixture was vortexed and DMSO (120μl) added drop-wise until the two layers became miscible. The precipitate formed was assumed to be salt and the reaction was left overnight at RT.

The product was concentrated to remove the acetonitrile before being desalted through an Illustra NAP™ column, lyophilised and re-dissolved in pure water.

The mass spectrum data of RNA1_D gave the expected mass of [M+H] = 8586.68 (Appendix Figure 5.9).

2.5.2.9 – Preparation of RNA1_F from RNA1_E and Compound 11



To a solution of RNA1_E (0.9μg in 66.7μl water) in NaHCO₃ (90μl, 0.025M, pH~8) was added a solution of compound **11** (1mg, 4.42x10⁻⁶mol) in DNA grade acetonitrile (210μl). The mixture was vortexed and DMSO (80μl) was added drop-wise until the two layers were miscible. The precipitate formed was assumed to be salt and the reaction was left overnight at RT.

The product was concentrated to remove the acetonitrile before being desalted through an Illustra NAP™ 25 column, lyophilised and re-dissolved in pure water. The mass spectrum data of RNA1_E gave the expected mass of [M+H] = 11683.65 (Appendix Figure 5.9).

2.5.3 – Click Chemistry of 2'OMe RNA Sequences

To investigate whether click chemistry reactions were possible between two 2'OMe RNA strands, reaction conditions to produce products RNA3-RNA6, RNA11 are to be investigated (Figure 2.37).

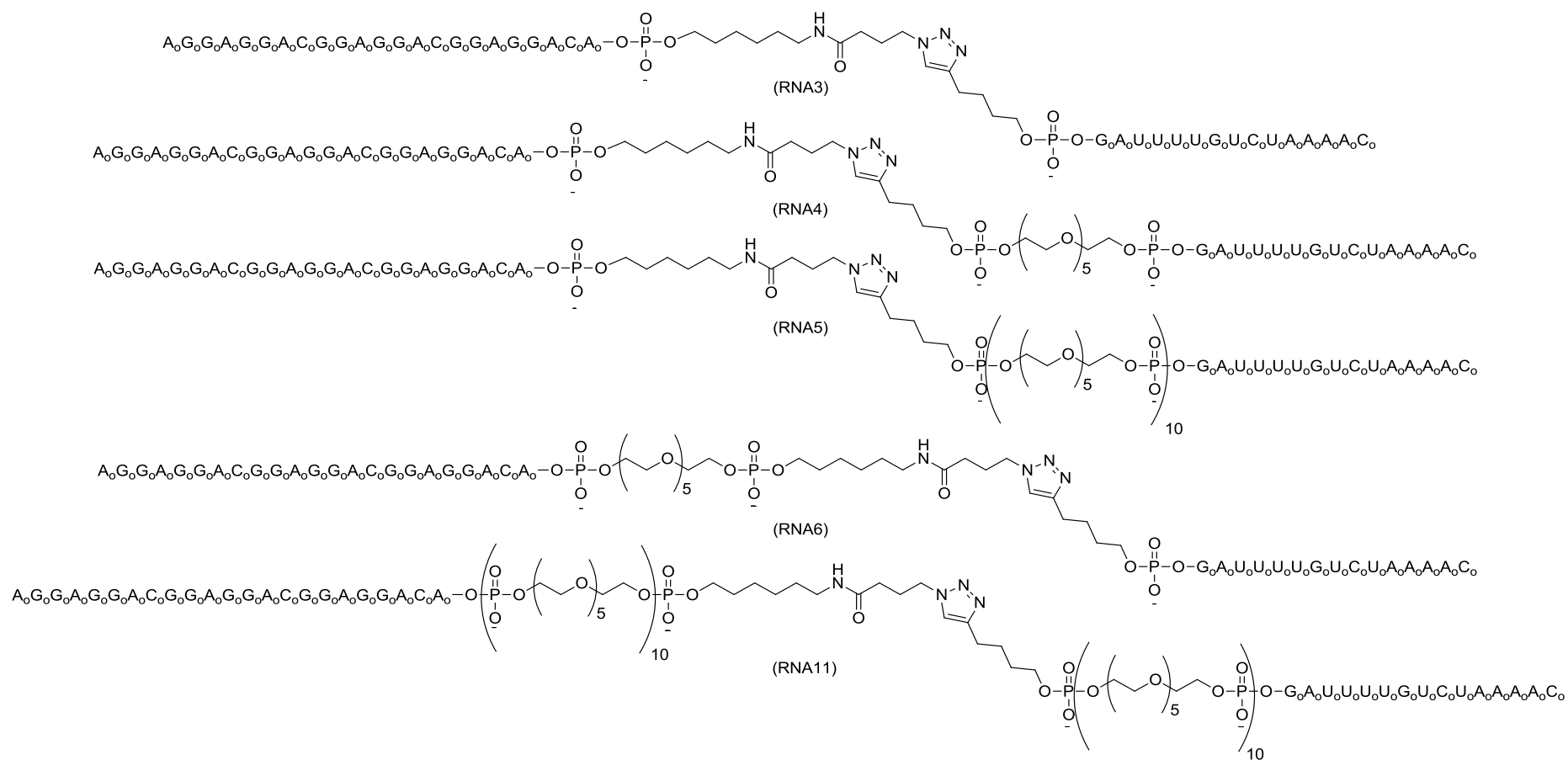


Figure 2.58. Products formed from Cu-catalysed Huisgen [3+2] cycloaddition reactions between RNA1 and RNA2 series strands.

2.5.3.1 – Preparation of RNA6 Tripartite Structure

RNA6 (Figure 2.37) was prepared from RNA1_D and RNA2_A by click chemistry. CuBr (4.4mg, 2.16×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 307 μ l, 0.1M 3), vortexed until olive green in colour and then 1 μ l diluted down with DMSO to 8.62mM. A further 1 μ l taken and diluted down with DMSO to 8.62×10^{-4} mol to investigate whether click chemistry would occur with 100eq or 10eq of copper. RNA1_D (5 μ l, 17.2 μ M) was mixed with RNA2_A (5 μ l, 17.2 μ M) before adding Cu-TBTA (1 μ l, 8.62mM solution, 100eq), with a second set of reaction being treated with Cu-TBTA (1 μ l, 8.62×10^{-4} mol solution, 10eq). The solutions were vortexed for one minute to ensure thorough mixing and then heated to 40°C at 750rpm. The reactions were monitored by taking 3 μ l samples at 15 minute, 30 minute and 1 hour time points, placing the samples in a -20°C freezer to stop the reaction.

The samples were desalted (MF-Millipore™ Membrane, 20mins) then analysed by gel electrophoresis by running a 10% denaturing PAGE at 100V for 1 hour. The gel was stained with SYBR™ gold and image captured on a VersaDoc™ imager (Section 2.4.4.1, Figure 2.14).

2.5.3.2 – Preparation of RNA11 Tripartite Structures

RNA11 (Figure 2.37) was prepared from RNA1_F and RNA2_C by click chemistry. CuBr (5.3mg, 3.70×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 369.5 μ l, 0.1M), vortexed until olive green in colour and then 1 μ l diluted down with DMSO to 8.62mM. RNA1_F (5 μ l, 17.2 μ M) was mixed with RNA2_C (5 μ l, 17.2 μ M) before adding Cu-TBTA (1 μ l, 8.62mM, 100eq). The solution was vortexed for one minute to ensure thorough mixing and then heated to 40°C at 750rpm. The reaction was monitored by taking 3 μ l

samples at 15 minutes, 30 minutes, 1 hour time points, placing the samples in a -20°C freezer to stop the reaction.

Controls were run alongside this reaction with only the 30 minute time point taken for the controls. The controls were:-

- 1) RNA1_F (17.2µM, 5ul) and water (5µl) with and without Cu-TBTA.
- 2) RNA2_C (17.2µM, 5ul) and water (5µl) with and without Cu-TBTA.
- 3) RNA1_F (17.2µM, 5ul) and RNA2_C (17.2µM, 5ul) without Cu-TBTA.

The samples were desalted (MF-Millipore™ Membrane, 20mins) then analysed by gel electrophoresis by running a 10% denaturing PAGE at 100V for 1 hour. The gel was stained with SYBR™ gold and image captured on a VersaDoc™ imager (Section 2.4.4.1, Figure 2.16).

2.5.3.3 – All 2'OMe RNA Click Chemistry Reactions

To investigate whether 2'OMe RNA click chemistry reactions were possible for all tripartite constructs RNA3-RNA11 were prepared (Figure 2.38). CuBr (2mg, 1.39×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 139.4µl, 0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. Three tubes of RNA2_A (5µl, 17.2µM), three tube of RNA2_B (5µl, 17.2µM) and three tube of RNA2_C (5µl, 17.2µM) were prepared. RNA1_B (5µl, 17.2µM) was then added to the first tube for each of the RNA2 strands. The second tube of each of the RNA2 strands was treated with RNA1_D (5µl, 17.2µM) and the last three tubes treated with RNA1_F (5µl, 17.2µM). The reactions were then treated with Cu-TBTA (1ul, 8.62mM, 100eq), vortexed for one minute and heated to 40°C at 750rpm for 15 minutes. The

samples were desalted (MF-Millipore™ Membrane, 20mins) then analysed by 10% denaturing PAGE run at 100V for 1 hour. The gels were stained with SYBR™ gold and image captured on a VersaDoc™ imager. Controls of the relevant starting materials were also included on the gels (*Section 2.4.5*, Figure 2.18).

2.5.4 – Click Chemistry Scale Up Reactions

2.5.4.1 – Trial Purification Condition of Scale Up Click

RNA6 (Figure 2.38) was prepared from RNA1_D and RNA2_A by click chemistry. CuBr (3.2mg, 2.23×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 223µl, 0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. RNA1_D (5µl, 17.2µM) was mixed with RNA2_A (5µl, 17.2µM) before adding Cu-TBTA (1µl, 8.62mM, 100eq). Nine identical reactions were carried out at the same time to enable isolation of a reasonable amount of material. The solutions were vortexed for one minute to ensure thorough mixing and then heated to 40°C at 750rpm for 15 minutes. The nine reactions were combined and then split into three to trial purification techniques.

The first batch was purified by ion exchange (10mM Hepes, 20mM NaCl to 10mM Hepes, 1M NaCl) using one injection (Appendix Figure 5.15). The second batch of sample was purified by 10% denaturing PAGE run at 100V for 1 hour. The gel was stained with SYBR™ gold and image captured on a VersaDoc™ imager, the relevant bands cut out of the gel and then treated with RNA elution buffer (350µl) overnight at 5°C.

After the incubation the samples were heated to 30°C for 10 minutes to re-dissolve the crashed out SDS before transferring the supernatant to a new tube. To the supernatant ethanol (1.05ml, 3x vol of RNA elution buffer) was added, vortexed and then centrifuged for 30 minutes at 13,400rpm. The supernatant was then removed and the pelleted RNA washed with ethanol (250µl) before being centrifuged for 15 minutes at 13,400rpm. The supernatant was once again removed and the pellet dried and re-suspended in water.

The purification using gel electrophoresis gave the greatest product recovery and so the final batch of sample was also purified by gels.

2.5.4.2 – RNA3 Scale Up

RNA3 (Figure 2.38) was prepared from RNA1_B and RNA2_A by click chemistry. CuBr (1.6mg, 1.12×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 111.5µl, 0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. RNA1_B (60µl, 17.2µM) and RNA2_A (60µl, 17.2µM) were vortexed before carrying out the experiment as described in the general procedure click chemistry scale up. Purified products were then run against starting materials to confirm purity (*Section 2.4.6*, Figure 2.19). Compound RNA3 was dissolved in ddH₂O to give a 2µM solution ready for splicing reactions.

2.5.4.3 – RNA4 Scale Up

RNA4 (Figure 2.38) was prepared from RNA1_B and RNA2_B by click chemistry. CuBr (2.6mg, 1.81×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 181.3µl, 0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. RNA1_B (60µl, 17.2µM) and

RNA2_B (60µl, 17.2µM) were vortexed before carrying out the experiment as described in the general procedure click chemistry scale up. Purified products were then run against starting materials to confirm purity (*Section 2.4.6*, Figure 2.19). Compound RNA4 was dissolved in ddH₂O to give a 2µM solution ready for splicing reactions.

2.5.4.4 – RNA5 Scale Up

RNA5 (Figure 2.38) was prepared from RNA1_B and RNA2_C by click chemistry. CuBr (1.7mg, 1.19x10⁻⁵mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 118.5µl, 0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. RNA1_B (60µl, 17.2µM) and RNA2_C (60µl, 17.2µM) were vortexed before carrying out the experiment as described in the general procedure click chemistry scale up. Purified products were then run against starting materials to confirm purity (*Section 2.4.6*, Figure 2.19). Compound RNA5 was dissolved in ddH₂O to give a 2µM solution ready for splicing reactions.

2.5.4.5 – RNA6 Scale Up

RNA6 (Figure 2.38) was prepared from RNA1_D and RNA2_A by click chemistry. CuBr (3.8mg, 2.65x10⁻⁵mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 264.9µl, 0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. RNA1_D (60µl, 17.2µM) and RNA2_A (60µl, 34.4µM) were vortexed before carrying out the experiment as described in the general procedure click chemistry scale up. Purified products were then run against starting materials to confirm purity (*Section 2.4.6*, Figure 2.19). Compound RNA6 was dissolved in ddH₂O to give a 2µM solution ready for splicing reactions.

2.5.4.6 – RNA7 Scale Up

RNA7 (Figure 2.38) was prepared from RNA1_D and RNA2_B by click chemistry. CuBr (1.6mg, 1.12×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 111.5µl, 0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. RNA1_D (60µl, 17.2µM) and RNA2_B (60µl, 34.4µM) were vortexed before carrying out the experiment as described in the general procedure click chemistry scale up. Purified products were then run against starting materials to confirm purity (*Section 2.4.6*, Figure 2.19). Compound RNA7 was dissolved in ddH₂O to give a 2µM solution ready for splicing reactions.

2.5.4.7 – RNA8 Scale Up

RNA8 (Figure 2.38) was prepared from RNA1_D and RNA2_C by click chemistry. CuBr (1.5mg, 1.05×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 104.6µl, 0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. RNA1_D (60µl, 17.2µM) and RNA2_C (60µl, 34.4µM) were vortexed before carrying out the experiment as described in the general procedure click chemistry scale up. Purified products were then run against starting materials to confirm purity (*Section 2.4.6*, Figure 2.19). Compound RNA8 was dissolved in ddH₂O to give a 2µM solution ready for splicing reactions.

2.5.4.8 – RNA9 Scale Up

RNA9 (Figure 2.38) was prepared from RNA1_F and RNA2_A by click chemistry. CuBr (2.9mg, 2.02×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 202.2µl, 0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. RNA1_F (60µl, 17.2µM) and

RNA2_A (60µl, 34.4µM) were vortexed before carrying out the experiment as described in the general procedure click chemistry scale up. Purified products were then run against starting materials to confirm purity (*Section 2.4.6*, Figure 2.19). Compound RNA9 was dissolved in ddH₂O to give a 2µM solution ready for splicing reactions.

2.5.4.9 – RNA10 Scale Up

RNA10 (Figure 2.38) was prepared from RNA1_F and RNA2_B by click chemistry. CuBr (1.5mg, 1.05×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 104.6µl, 0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. RNA1_F (60µl, 17.2µM) and RNA2_B (60µl, 34.4µM) were vortexed before carrying out the experiment as described in the general procedure click chemistry scale up. Purified products were then run against starting materials to confirm purity (*Section 2.4.6*, Figure 2.19). Compound RNA10 was dissolved in ddH₂O to give a 2µM solution ready for splicing reactions.

2.5.4.10 – RNA11 Scale Up

RNA11 (Figure 2.38) was prepared from RNA1_F and RNA2_C by click chemistry. CuBr (0.8mg, 5.58×10^{-6} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 55.8µl, 0.1M), vortexed until olive green in colour and then 1µl diluted down with DMSO to 8.62mM. RNA1_F (60µl, 17.2µM) and RNA2_C (60µl, 17.2µM) were vortexed before carrying out the experiment as described in the general procedure click chemistry scale up. Purified products were then run against starting materials to confirm purity (*Section 2.4.6*, Figure 2.19). Compound RNA11 was dissolved in ddH₂O to give a 2µM solution ready for splicing reactions.

2.5.5 – 2'OMe Phosphorothioate (pS) RNA

The phosphorothioate series were prepared with PADS oxidiser on an ABI 394 synthesiser (Table 2.18) and were prepared with the same conjugation groups in order to undergo click chemistry (Table 2.19 and Schemes 2.6 and 2.7).

Table 2.18. Phosphorothioate products pS RNA1_A, pS RNA1_C, pS RNA1_E, pS RNA2_A-C prepared using the ABI synthesizer (Section 2.5.5.4). pS RNA1_B, pS RNA1_D and pS RNA1_E were formed from the NHS coupling reaction of pS RNA1_A, pS RNA1_C and pS RNA1_E with compound 11 (Section 2.5.5.6). W, X, Y, and Z are shown below, _o represents 2'OMe RNA and _s represents pS.

Sequence Name	Sequence 5' – 3'
pS RNA2_A	Y _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o
pS RNA2_B	Y _s X _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o
pS RNA2_C	Y _s X _s X _s X _s X _s X _s X _s X _s X _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} C _o
pS RNA1_A	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} Z
pS RNA1_B	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} W
pS RNA1_C	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s Z
pS RNA1_D	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s W
pS RNA1_E	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s X _s X _s X _s X _s X _s X _s X _s X _s Z
pS RNA1_F	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s X _s X _s X _s X _s X _s X _s X _s X _s W

W =

X =

Y =

Z =

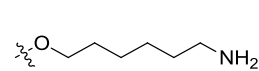
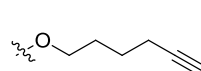
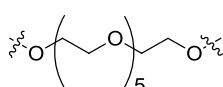
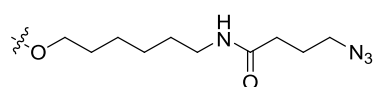
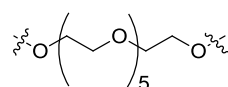


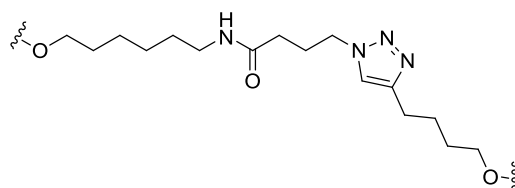
Table 2.19. Products formed from click chemistry of starting materials. X and Z structure are given below, _o represents 2'OMe RNA and _s represents pS.

Sequence Name	Sequence 5' – 3'	Starting Materials
pS RNA3	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _o sG _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} Z _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _o sC _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o	pS RNA1_B pS RNA2_A
pS RNA4	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _o sG _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s Z _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o	pS RNA1_B pS RNA2_B
pS RNA5	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _o sG _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s X _s X _s X _s X _s X _s X _s X _s X _s Z _s G _o sA _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o	pS RNA1_B pS RNA2_C
pS RNA6	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _o sG _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s Z _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o	pS RNA1_D pS RNA2_A
pS RNA7	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _o sG _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s Z _s X _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _o osU _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o	pS RNA1_D pS RNA2_B
pS RNA8	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _o sG _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s Z _s X _s X _s X _s X _s X _s X _s X _s X _s X _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o	pS RNA1_D pS RNA2_C
pS RNA9	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _o sG _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s X _s X _s X _s X _s X _s X _s X _s X _s Z _s G _o sA _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o	pS RNA1_F pS RNA2_A
pS RNA10	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _o sG _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s X _s X _s X _s X _s X _s X _s X _s X _s Z _s X _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o	pS RNA1_F pS RNA2_B
pS RNA11	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _o sG _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s X _s X _s X _s X _s X _s X _s X _s X _s Z _s X _s X _s X _s X _s X _s X _s X _s X _s X _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} C _o	pS RNA1_F pS RNA2_C

X =

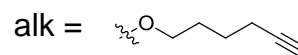


Z =



2.5.5.1 – Synthesis of pS DNA 3

5' alk_sG_sA_sT_sT_sT_sT_sG_sT_sC_sT_sA_sA_sA_sA_sC 3'



pS-DNA 3 was prepared by the DMT- and β -(cyanoethyl) phosphoramidite method on 1 μ mol CPG supports using an Applied Biosystems 394 machine. Bis (phenyl acetyl) disulfide (PADS) (0.2M in 1:1 acetonitrile/pyridine) which had been “aged” for two days, was used as the oxidiser¹⁶⁰. Oxidising times were extended to 5 second additions with 2 minutes pausing intervals repeated three times. Coupling times for all phosphoramidites were 11 minutes.

After automated synthesis the pS DNA 3 was treated with 1:1 triethylamine/acetonitrile for 2 hours before being cleaved and deprotected from the solid support by aminolysis (35% aqueous ammonia) incubating the resin at 55°C for 16 hours. The aqueous ammonia was removed by evaporation and the product desalted using an Illustra NAPTM 25 column before pS DNA3 was confirmed by MALDI-TOF ([M+H] = 4967.96) (Section 2.4.7, Figure 2.20).

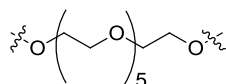
2.5.5.2 – Preparation of pS 2'OMe RNA Strands

pS-2'OMe RNA was prepared by the DMT- and β -(cyanoethyl) phosphoramidite method on 1 μ mol CPG supports using an Applied Biosystems 394 machine. Bis (phenyl acetyl) disulfide (PADS) (0.2M in 1:1 acetonitrile/pyridine), which had been “aged” for two days, was used as the oxidiser¹⁶⁰. Oxidising times were extended to 5 second additions with 2 minutes pausing intervals repeated three times. Coupling times for the RNA2 series phosphoramidites were 11 minutes with a double coupling protocol used for the RNA1 series with coupling times extended to 18 minutes.

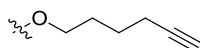
Table 2.20. *pS RNA2_A, pS RNA2_B, pS RNA2_C, pS RNA1_A, pS RNA1_C and pS RNA1_E were prepared by solid phase synthesis. X, Y and Z structure are given below, _o represents 2'OMe RNA and _s represents pS*

Sequence Name	Sequence – 5' to 3'
pS-RNA2_A	Y _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o
pS RNA2_B	Y _s X _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o
pS RNA2_C	Y _s X _s X _s X _s X _s X _s X _s X _s X _s X _s G _{os} A _{os} U _{os} U _{os} U _{os} U _{os} G _{os} U _{os} C _{os} U _{os} A _{os} A _{os} A _{os} A _{os} C _o
pS RNA1_A	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} Z
pS RNA1_C	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s Z
pS RNA1_E	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s X _s X _s X _s X _s X _s X _s X _s X _s Z

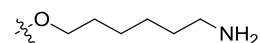
X =



Y =



Z =



After automated synthesis the strands were treated with 1:1 triethylamine/acetonitrile for 2 hours before being cleaved and deprotected by aminolysis (35% aqueous ammonia) incubating the resin at 55°C overnight. The aqueous ammonia was removed by evaporation and the products desalted using an Illustra NAP™ 25 column. Characterisation of pS strands were problematic as the samples showed no distinct mass by MALDI-TOF or accurate mass (Appendix Figures 5.12).

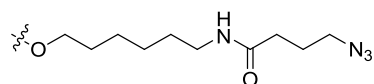
2.5.5.3 – Preparation of 2'OMe pS RNA Azides

Table 2.21. RNA products prepared from the NHS reaction with compound **11**.

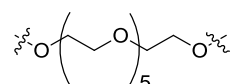
W, X and Z structure are given below, _o represents 2'OMe RNA and _s represents pS

Sequence Name	Sequence 5' – 3'	Starting Material
RNA1_B	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _o sA _{os} G _{os} G _{os} A _{os} C _{os} A _{os} W	RNA1_A
RNA1_D	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _o sA _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s W	RNA1_C
RNA1_F	A _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _{os} A _{os} G _{os} G _{os} A _{os} C _{os} G _{os} G _o sA _{os} G _{os} G _{os} A _{os} C _{os} A _{os} X _s X _s X _s X _s X _s X _s X _s X _s X _s W	RNA1_E

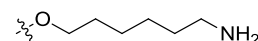
W =



X =



Z =



pS RNA1_A, pS RNA1_C and pS RNA1_E (200µl) were evaporated to dryness, in separate tube, before each being re-dissolved in aqueous NaHCO₃ (180µl, 0.025M, pH~8). A solution of compound **11** (9mg, 1.33 x 10⁻⁵ mol) in DNA grade acetonitrile (1.26ml) was prepared before adding 420µl to each of the RNA1 solutions. The mixture was vortexed and left overnight at RT. The product was concentrated to remove the acetonitrile and desalted through a Illustra NAPTM 25 column.

2.5.6 – All pS Click Chemistry Reactions

To investigate whether pS 2'OMe RNA click chemistry reactions were possible for all tripartite constructs pS RNA3-pS RNA11 were prepared (Figure 2.39).

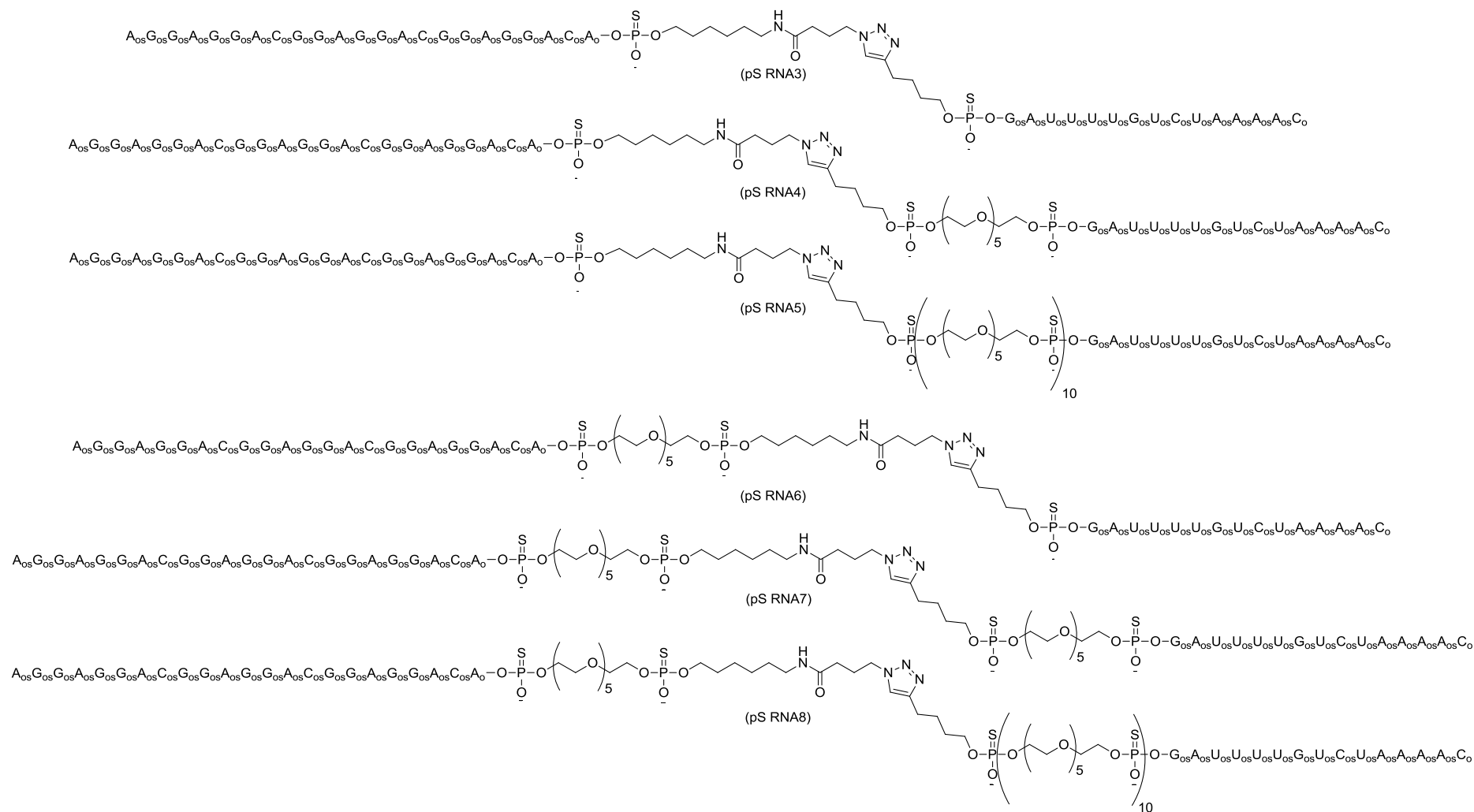


Figure 2.39. Products of the Cu-catalysed Huisgen [3+2] cycloaddition bioconjugation reactions for pS 2'OMe RNA.

CuBr (2.2mg, 1.53×10^{-5} mol) was added to solution of TBTA in DMSO/t-butanol (3:1, 153.36 μ l, 0.1M), vortexed until olive green in colour and then 1 μ l diluted down with DMSO to 8.62mM.

Three tubes of pS RNA2_A (5 μ l, 17.2 μ M), three tubes of pS RNA2_B (5 μ l, 17.2 μ M) and three tubes of pS RNA2_C (5 μ l, 17.2 μ M) were prepared. pS RNA1_B (5 μ l, 17.2 μ M) was then added to one of each the pS RNA2 strands. The same was done for pS RNA1_D and pS RNA1_F to test all possible click chemistry combination. The reactions were then treated with CuTBTA (1 μ l, 8.62mM, 100eq), vortexed and then heated to 40°C at 750rpm for 15 minutes.

The samples were desalted (MF-Millipore™ Membrane, 20 minutes) before being organised so that each reaction had the starting materials on the same gel to ensure product was being formed. Three 20% denaturing polyacrylamide gels were run at the same time at 4W for 3 hours. The gels were stained with SYBR™ gold and image captured on a VersaDoc™ imager (Section 2.4.10, Figure 2.25).

2.5.7 – pS Scale Up Click

2.5.7.1 – pS RNA6 Scale Up

pS RNA6 (Figure 2.39) was prepared from pS RNA1_D and pS RNA2_A by click chemistry. CuBr (1.5mg, 1.05×10^{-5} mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 104.6 μ l, 0.1M), vortexed until olive green in colour and then 1 μ l diluted down with DMSO to 8.62mM. pS RNA1_D (60 μ l, 17.2 μ M) and pS RNA2_A (60 μ l, 51.6 μ M) were vortexed and then the experiment carried out as in the general click chemistry scale up procedure.

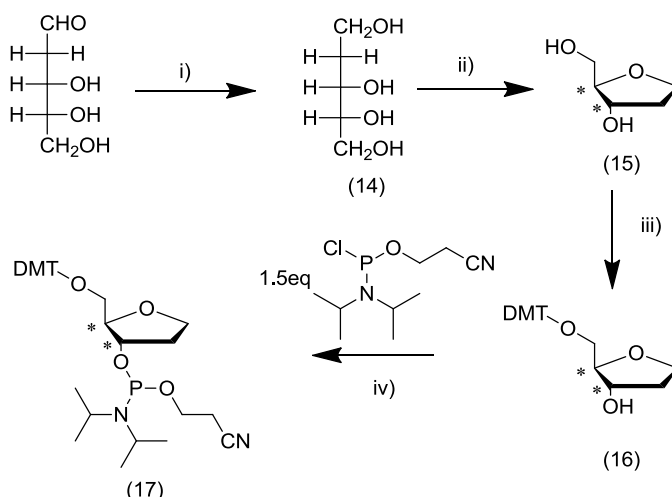
Four 20% denaturing polyacrylamide gels were run at 4W for 3 hours to purify the product with the work up the same as the general click chemistry scale up procedure. Purified products were then run against starting materials to confirm purity (*Section 2.4.12*, Figure 2.26). Compound pS RNA6 was dissolved in ddH₂O to give a 2μM solution ready for splicing reactions.

2.5.7.2 – pS RNA7 Scale Up

pS RNA7 (Figure 2.39) was prepared from pS RNA1_D and pS RNA2_B by click chemistry. CuBr (1.6mg, 1.12x10⁻⁵mol) was added to a solution of TBTA in DMSO/t-butanol (3:1, 111.5μl, 0.1M), vortexed until olive green in colour and then 1μl diluted down with DMSO to 8.62mM. pS RNA1_D (60μl, 17.2uM) and pS RNA2_B (60μl, 51.6μM) were vortexed and then the experiment carried out as in the general click chemistry scale up procedure.

pS RNA7 products were desalted, heated to 80°C and split over four 20% denaturing polyacrylamide gels and ran at 4W for 4¹/₂ hours. Two bands were cut out and worked up individually as described in the general click chemistry scale up procedure. The two bands were analysed by 20% denaturing polyacrylamide gel to ensure which band was product and which starting material. The samples were heated to 80°C before loading to ensure the secondary structure was fully melted. Purified products were then run against starting materials to confirm purity (*Section 2.4.12*, Figure 2.26). Compound pS RNA7 was dissolved in ddH₂O to give a 2μM solution ready for splicing reactions.

2.5.8 – Synthesis of Abasic Strands



Scheme 2.8. Highlights the synthetic route to make the abasic phosphoramidite **17**. * represent chiral centres. i) NaBH₄ in 0.1mM NaOH. ii) HCl 2N at 95°C for 3 days. iii) 1eq DMT-Cl, 0.1eq DMAP in pyridine for 16 hours. iv) 1.3eq DIPEA in DCM

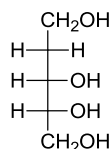
Abasic strands were prepared according to the protocol described by R. Eritja *et al* (Scheme 2.8)⁶. The phosphoramidite **17** was used in solid phase synthesis on an ABI 394 synthesiser to make the relevant strands (Table 2.22).

Table 2.22. RNA14 and RNA1_G prepared by solid phase synthesis (Section 2.5.8.4). Compound RNA 1_H formed from the NHS coupling reaction of RNA1_G with compound **11** (Section 2.5.8.5). W, X and Z are shown below and _o represents 2'OMe RNA.

Sequence Name	Sequence 5' – 3'
RNA14	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XXXXXX G _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o C _o
RNA1_G	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XXXZ
RNA1_H	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o XXXW

W = X = Z =

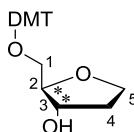
2.5.8.1 – Preparation of (2R,3S)-pentane-1,2,3,5-tetraol (**14**)



(14)

Compound **14** was synthesised according to a protocol described by R. Eritja *et al*⁶. 2-Deoxy-D-ribose (2.1g, 15mmol) was added to a solution of NaBH₄ (215mg, 5.68mmol) in 0.1M sodium hydroxide solution (11ml) and left overnight stirring at room temperature. Dowex 1 x 8 (100-200dry mesh size) was activated by washing the resin five times with 1M sodium hydroxide (200ml), allowing the resin to settle before removing the supernatant between washings. The resin was washed with water (eight 300ml washings) until the pH was approximately pH 7. At the same time Dowex 50W x 8 (100-200 dry mesh size) was activated by washing the resin five times with 1M HCl (200ml), allowing the resin to settle before removing the supernatant. The resin was washed with water (eight 300ml washings) until the pH was approximately pH7. The two resin were then stored overnight. The product was acidified to pH 4 using 50% acetic acid to quench the reaction. The reaction mixture was passed through a column of Dowex 50W x 8 resin (25ml resin). The column was flushed with five column volumes to ensure the product had been eluted. The eluted product went into a conical flask containing 25ml Dowex 1 x 8 which was mixed and then transferred into a column to elute the product. Elution of the product was achieved using five column volumes of water followed by removal of the water on a rotary evaporator. The green oil product (compound **14**) was then used crude in the next reaction.

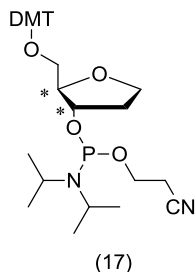
2.5.8.2 – Preparation of (2R,3S)-2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)tetrahydrofuran-3-ol (**16**)



(16)

Compound **16** was synthesised according to a protocol described by R. Eritja *et al*⁶. 2M HCl (20ml) was added to compound **14** and the reaction mixture was heated to for 5 days. After 5 days the HCl was removed on a rotary evaporator and the brown oil (Compound **15**). Compound **15** (2.20g, 18.69×10^{-3} mol) was treated with anhydrous pyridine (12.5ml) and then concentrated *in vacuo*. This process was repeated three times to ensure the HCl from the previous step had been neutralised. Compound **15** (2.20g, 18.69×10^{-3} mol) in pyridine (10ml) was then transferred to a round bottom flask with molecular sieves (4Å) and left stirring under nitrogen for 7 hours. A solution of DMAP (15.2mg, 1.25×10^{-3} mol) in anhydrous pyridine (2ml) was added to the reaction followed by the addition of a suspension of DMT-Cl (4.23g, 12.46×10^{-3} mol) in anhydrous pyridine (7ml) before leaving stirring under nitrogen for 16 hours. Column chromatography (SiO₂, 20% ethyl acetate/petroleum ether → 50% ethyl acetate afforded compound **16** as a yellow gum in 15 % (1.136 g); ES/MS: m/z 443 (+Na), ¹H NMR (400 MHz; CDCl₃) δ: 1.76-1.84 (m, 1 H, C4) 2.01-2.12 (m, 1 H, C4) 2.97-3.03 (m, 1 H, C1) 3.13-3.19 (m, 1 H, C1) 3.77-3.83 (m, 1 H, C2), 3.71 (s, 6 H, DMT) 3.87-3.92 (m, 2 H, C5) 4.18-4.24 (m, 1 H, C3) 6.72-6.77 (m, 4 H, DMT) 7.09-7.15 (m, 1 H, DMT) 7.17-7.27 (m, 6 H, DMT) 7.33-7.38 (m, 2 H, DMT). This assignment was confirmed by COSY and literature precedent⁶.

2.5.8.3 – Preparation of 2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)tetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (**17**)



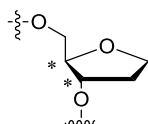
Compound **17** was synthesised according to a protocol described by R. Eritja *et al*⁶. A solution of compound **16** (300mg, 7.13×10^{-4} mol), diisopropylethylamine (161.5 μ l, 9.27×10^{-4} mol) in degassed anhydrous DCM (2ml) was added to a flame dried, argon filled flask. The flask was briefly evacuated and re-filled with argon before the addition of a solution of 2-cyanoethyl N,N-diisopropylchlorophosphoramidite (238.7 μ l, 1.07×10^{-3} mol) in dry DCM (2ml). The reaction mixture was stirred at RT for 4 hrs. The product was purified by column chromatography using neutralised SiO₂ (neutralised with pyridine and heated in an oven overnight at 120°C; eluent 50% dry ethyl acetate/ dry hexane with 0.1% triethylamine). Dried 100ml round bottom flasks were used to collect fractions, collecting five 50ml fractions. The solvent was removed by direct evaporation on a high vacuum line fitted with a cold trap affording compound **17** as a colourless oil in 74% (154.3mg); ¹H NMR (400 MHz; CDCl₃) δ : 0.95-1.19 (m, 13 H) 1.81-1.93 (m, 1 H) 1.94-2.08 (m, 1 H) 2.26-2.32 (m, 1 H) 2.97-3.12 (m, 2 H) 3.39-3.59 (m, 3 H) 3.65 (s, 7 H) 3.79-3.99 (m, 3 H) 4.26-4.37 (m, 1 H) 6.67-6.76 (m, 4 H) 7.04-7.21 (m, 4 H) 7.22-7.28 (m, 3 H) 7.34-7.39 (m, 2 H), ³¹P NMR (161MHz; CDCl₃) δ : 148.20 (s, 1 P) 147.87 (s, 1 P). This assignment was confirmed literature precedent⁶.

2.5.8.4 – Preparation of 2'OMe RNA Abasic Strands

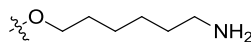
Table 2.23. RNA14 and RNA1_G prepared by solid phase synthesis. X and Z are shown below, _o represents 2'OMe RNA and * represents a chiral centre.

Sequence Name	Sequence – 5' to 3'
RNA14	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o X XXXXXG _o A _o U _o U _o U _o U _o G _o U _o C _o U _o A _o A _o A _o C _o
RNA1_G	A _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o G _o G _o A _o G _o G _o A _o C _o A _o X XXZ

X =



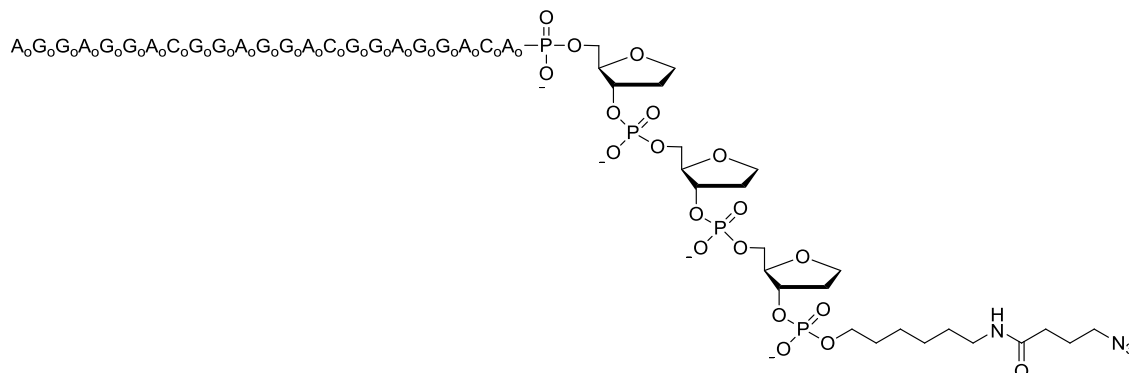
Z =



2'OMe RNAs were prepared by the DMT- and β -(cyanoethyl) phosphoramidite method on 1 μ mol CPG supports using an Applied Biosystems 394 machine. Double coupling procedures were used for both sequences with coupling times extended to 18 minutes. Upon completion of the automated synthesis the sequences were cleaved from the solid support and deprotected by aminolysis (35% aqueous ammonia) at 55°C overnight. The aqueous ammonia was removed by evaporation and the samples desalted through an Illustra NAPTM 25 column, lyophilised and re-dissolved in ddH₂O.

RNA14 (180 μ l, 17.2 μ M) was purified by gel electrophoresis by adding to loading dye (72 μ l), heating to 80°C before loading on four 20% denaturing polyacrylamide gels (Section 2.4.3.2, Figure 2.11). The gels were run at 4W for 4 hours and the product worked up as highlighted in the general click chemistry protocol. Compound RNA14 was dissolved up in ddH₂O to give a 2 μ M solution ready for splicing reactions.

2.5.8.5 – Preparation of RNA1_H Conjugate



RNA1_G (137.4μM, 250μl) was evaporated to dryness before being re-dissolved in aqueous NaHCO₃ (180μl, 0.025M, pH~8). Compound **11** (2mg, 8.84x10⁻⁶mol) was dissolved in DNA grade acetonitrile (420μl) before adding to the RNA15 solution. The mixture was vortexed, a precipitate formed and the reaction was left overnight at RT. The product was concentrated to remove the acetonitrile, desalted through an Illustra NAP™ 25 column and then lyophilised. The product was re-dissolved in 200μl water.

2.5.9 – RNA15 Scale Up

CuBr (2mg, 1.39x10⁻⁵ mol) was dissolved in a TBTA solution (139.4μl, 0.1M 3:1 DMSO/t-butanol), vortexed until olive green in colour before 1μl was taken and diluted in DMSO to 8.62mM. RNA1_H (60μl, 17.2uM) and RNA2_A (60μl, 51.6μM) were vortexed and then the experiment carried out as in the general click chemistry scale up procedure. Four 20% denaturing polyacrylamide gels were run at 4W for 3¹/₂ hours to purify the product with the work up the same as the general click chemistry scale up procedure. Compound RNA15 (Figure 2.40) was dissolved in ddH₂O to give a 2μM solution ready for splicing reactions.

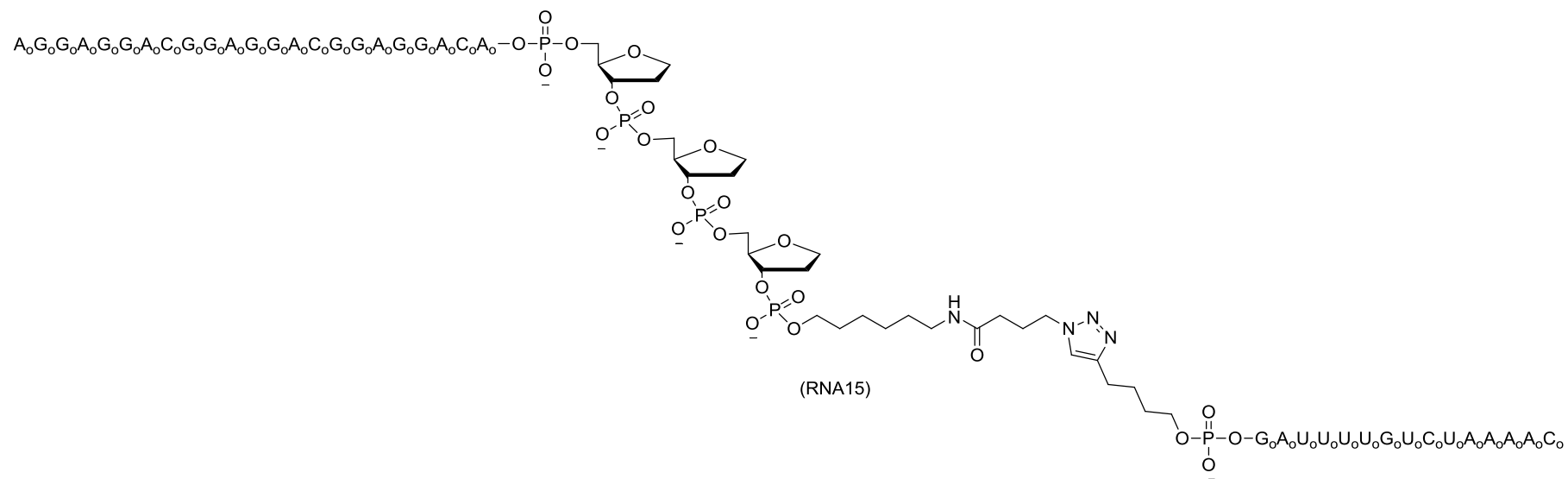
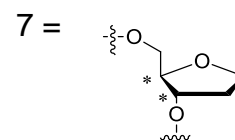


Figure 2.40. Product of the Cu-catalysed Huisgen [3+2] cycloaddition bioconjugation reaction forming RNA15

2.5.10 – Solid Phase Synthesis of RNA16

5'A_sG_sG_sA_sG_sGACGGAGGACGGAGGACA777777G_oA_oU_{os}U_{os}U_{os}U_{os}G_{os}U_oC_oU_o

A_oA_{os}A_{os}A_{os}C_o



The mixed RNA strand was prepared by the DMT- and β-(cyanoethyl) phosphoramidite method on 1 μmol CPG supports using an Applied Biosystems 394 machine. Bis (phenyl acetyl) disulfide (PADS) (0.2M in 1:1 acetonitrile/pyridine), which had been “aged” for two days¹⁶⁰, was used as the oxidiser with oxidising times being extended to 5 second additions with 2 minute pauses repeated three times. Coupling times for the bases in bold were 11 minutes. Double coupling procedures were used for the rest of the sequences with coupling times extended to 18 minutes. The above strand was synthesised by splitting the strand into 6 separate sequences due to the need to switch oxidisers and bases at various points during the synthesis:-

- Sequence 1 - A_{os}A_{os}A_{os}C_o - used 2'OMe RNA and the 0.2M PADS oxidiser.
- Sequence 2 – U_oC_oU_oA_oA_o - used 2'OMe RNA and the 0.02M iodine oxidiser.
- Sequence 3 - U_{os}U_{os}U_{os}U_{os}G_{os}U_o – used 2'OMe RNA and the 0.2M PADS oxidiser.
- Sequence 4 - 777777G_oA_oU_o – used the 2'OMe RNA, the DNA abasic linker and the 0.02M iodine oxidiser.
- Sequence 5 – GACGGAGGACGGAGGACA – used TC RNA and the 0.02M iodine oxidiser.

- Sequence 6 - A_sG_sG_sA_sG_sG – used TC RNA and the 0.2M PADS oxidiser.

After automated synthesis the strand was washed twice with 20% diethylamine in acetonitrile (2.5ml, 3 minute washings), drying under argon between washings. The column was then washed three times with dry acetonitrile (2.5ml) and then washed twice with 1:1 triethylamine/acetonitrile (2.5ml) for 3 minutes before drying again. Three additions of acetonitrile (2.5ml) were then added again to the columns to wash them, drying the columns between washes. The resin was then removed from the column and treated with 1:1 ethylene diamine (EDA)/toluene (500µl) for 4 hours at 750rpm. The resin and solution were then transferred into filter tubes leaving the solution to drip slowly through. The filter tube was then transferred into new eppendorf tubes and centrifuged at 9,000rpm for 10 minutes before washing the resin with acetonitrile (500µl). The filter tube was centrifuged again at 9,000rpm for 10 minutes before washing the resin three times with acetonitrile (500µl), allowing the solution to drip through under gravity. The filter tube was then transferred to a new tube and centrifuged again at 9,000rpm for 10 minutes. The product was eluted using TE.1 solution (two 500µl elutions) and the resin dried by centrifuging at 9,000rpm for 10 minutes to remove all the TE.1 solution. The product was desalted through an Illustra NAPTM 25 column, lyophilised and re-dissolved in water. 450µl of the product was then purified by a 15% denaturing polyacrylamide gel (running 150µl per lane), using short wavelength UV light to determine where the bands were. The cut bands were placed in separate tubes, treated with RNA elution buffer (350µl per tube) and left at 5°C overnight. After incubation the samples were heated to 30°C for 10 minutes to re-dissolve the precipitated SDS before

transferring the supernatants to new tubes. Ethanol (1.05ml, 3x vol of RNA elution buffer) was added to the supernatant, vortexed and then centrifuged for 30 minutes at 13,400rpm. The supernatant was then removed and the pelleted RNA washed with ethanol (250µl) before being centrifuging again for 15 minutes at 13,400rpm. The supernatant was once again removed, the pellets dried and re-suspended in water (*Section 2.4.13*, Figure 2.49). Compound RNA16 was dissolved in ddH₂O to give a 2µM solution ready for splicing reactions.

Chapter 3 – Tripartite oligonucleotide splicing

3 – Introduction

Rectifying splicing defects is a well known strategy for preventing disease progression, with oligonucleotides having been used to modify splicing outcomes in cancer and other diseases^{161–166}. When the skipping of required exons occurs, re-introducing splicing to a particular exon requires either blocking of regulatory sequences within the pre-mRNA or increasing the number of positive acting factors to the skipped exon.

In this chapter the tripartite sequences prepared in chapter 2 are tested in SMN2 *in vitro* splicing assays to investigate what effect the bioconjugation group and linkers have on splicing.

3.1 – Aims

The bioconjugation used to conjugate the annealer and enhancer sequences was a Cu mediated click chemistry reaction as this provided a selective and fast conjugation. Since a bioconjugation group had not been tested in a splicing reaction before, the aim was to determine whether a bioconjugation group itself would be tolerated and whether linkers would be required to space the bioconjugation group away from the strands, allowing exon 7 inclusion levels to increase. Further aims of the splicing experiments were to determine whether increasing the affinity of proteins for the tripartite sequence (pS) would increase exon 7 inclusion levels and whether flexibility and/or the distance between the enhance and annealer within a tripartite sequence is beneficial for exon 7 inclusion levels.

3.2 – Blocking Regulatory Sequences to Modify Splicing

Outcomes

The main drawback of investigating the blocking of regulatory sequences is the cost of producing the large numbers of oligonucleotides that are required to determine the locations of silencer and enhancer sequences, which are difficult to predict. However, a blocking strategy using anti-sense oligonucleotides (ASO) has been shown to be a successful treatment, in Duchenne muscular dystrophy, cancer and other diseases^{161–166}. Oligonucleotides used in the blocking method are usually modified to increase both the affinity of the ASO for the pre-mRNA and resistance to degradation by nucleases. Structures of some of these modified bases are shown in Figure 3.1.

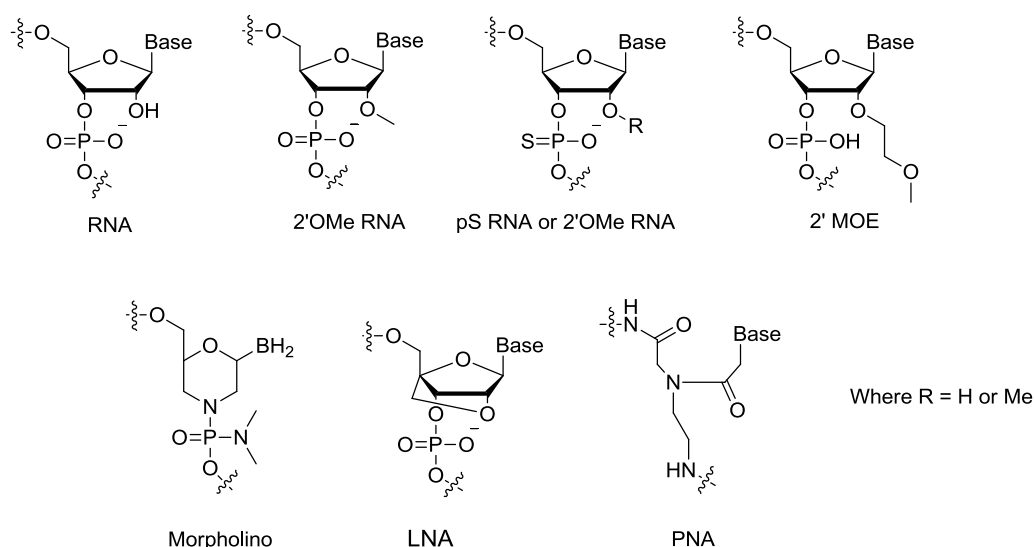


Figure 3.1. Structure of modified RNA strands used to increase pre-mRNA binding and nuclease resistance. Where MOE is 2' methoxy-ethyl ribose, LNA is locked nucleic acid and PNA is peptide nucleic acid.

Experiments using 2'OMe RNA ASO have been undertaken in *SMN2* to determine an ISS sequences in region +10 to +24 of intron 7¹⁶⁷. ASO binding

across this region caused a 5 fold increase in exon 7 inclusion levels *in vivo* at 50nM for *SMN2* fibroblast cells by preventing inhibitor proteins from binding ¹⁶⁷. This indicated that ASO binding across silencer sequences is an effective way to increase exon 7 inclusion levels. However, as silencer sequences are hard to predict, large numbers of ASO or extensive deletion studies (removing small amounts of introns from mini genes by endonuclease digestions and ligations) are required ^{167–169}. ASO screening of exons is preferable as exons are shorter than introns meaning less oligonucleotides would be required ¹³¹. In exons, as in introns, there are both enhancer and silencer sequences but due to the shorter length of the exons these sequences are much closer together. Annealing of ASOs within exons, to stimulate exon inclusion, will only be tolerated if the binding of enhancer proteins and compulsory spliceosomal factors are not affected. An ASO study carried out using 2' methoxy ethyl ribose (MOE) oligonucleotides (Figure 3.1) in *SMN2* showed that ASOs binding too close to the 3' SS (at positions +1 to +15 of exon 7) and ASOs binding over a tra2 β binding site (binding +16 to +30, +21 to +35 and +26 to +40 of exon 7) all inhibited splicing to that exon ¹³¹. Therefore screening of the annealing positions of ASOs is crucial in order to obtain the desired splicing outcome. The best exon based ASO produced gave rise to a ~1.5 fold increase in exon 7 inclusion levels in *in vitro* splicing at 100nM, when annealed to +7 to +21 of exon 7 ¹³¹. Another approach to reintroduce splicing to exon 7 in *SMN2* is to block the 5'SS of exon 8 using a modified U7 snRNA ¹⁷⁰. Blocking the splice site of exon 8 is not a problem in *SMN2* as functional SMN protein is encoded only in the first seven exons; therefore exon 8 can be skipped with no effect on SMN

production. This approach however may not be suitable for use in other disease systems.

3.3 – Introducing Positive Factors

Positive acting factors are important as they contain an interactive site that can bind or interact with enhancer proteins, which can increase splicing at a particular SS. This approach requires an annealer sequence which can be used in a similar way to ASO and bind to sites that silencer proteins would bind to, stopping them from binding. Two different positive acting factors have been investigated; the first being TOES (discussed *Section 2.1*) and the second being exon-specific splicing enhancement by small chimeric effectors (ESSENCE)¹⁷¹. ESSENCE uses a PNA annealer which is conjugated to ten RS repeats (Figure 3.2).

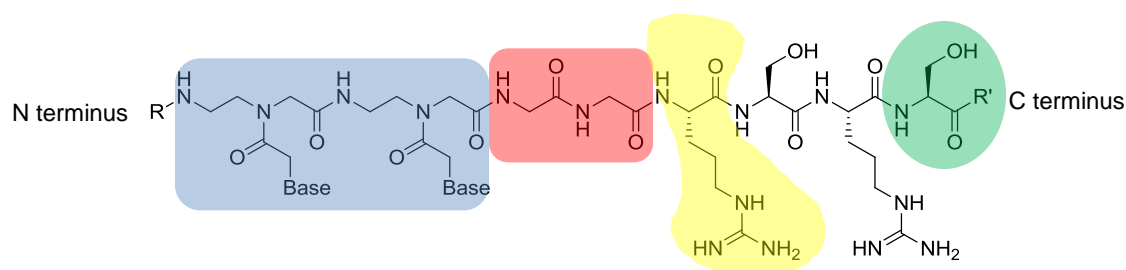


Figure 3.2. Structure of the key elements of an ESSENCE oligonucleotide. The blue box indicates PNA nucleotides, red indicates glycine linker, yellow indicates arginine peptide and green indicates serine peptide. The blue box also indicates the annealer region of the PNA-peptide and the structures in yellow and green make up the arginine-serine dipeptide. The digram shows 2 repeats of a RS domain the further eight are represented by R' [diagram adapted from Krainer]¹⁷¹.

When ten repeats of a non-natural RS domain are incorporated into SRSF-1, splicing can be increased at weak SSs even in the absence of normal SR protein SRSF-1. This is because a non-natural RS tail is able to form protein-protein interactions between SRSF-1 and U1-70K and therefore can be used to recruit SR proteins. PNA-peptide molecules, which have a non-natural RS tail, were constructed with the PNA annealer at the N terminal and the RS domain at the C terminus (Figure 3.2). PNA oligonucleotides were used rather than other RNA modifications, as the PNA backbone is neutral, creating no electrostatic repulsion between the pre-mRNA and the oligonucleotide. Other advantages of PNA are its resistance to endonucleases, the ease with which it can be synthesised by peptide-like chemistry and that fact that it crosses cell membranes and has been shown to cross the blood-brain barrier. Therefore, PNA oligomers provide an excellent backbone for rectifying splicing defects. Binding of the PNA-peptide sequence to +6 to +20 of exon 7 caused a 2.6 fold increase in exon 7 inclusion levels, whereas a PNA only control only increased exon 7 inclusion levels 1.7 fold at 400nM *in vitro*. These results reveal that the RS domain caused a 0.9 fold increase in exon 7 inclusion levels, meaning that ten repeats of the RS domain could efficiently recruit SR proteins to incorporate exon 7 into the mRNA.

Using ASO to block splicing signals in introns requires expensive screening to identify silencer sequences within the large span of introns. Blocking splicing signals within exons by ASO is preferred for screening (as exons are shorter) however the ASO effect on splice site selection can be further boosted by incorporating a binding/interactive site for enhancer proteins. PNA-peptide derivatives can be made by synthesising the PNA and peptide separately

before conjugating them to make the full length derivative, but TOES versions would require step wise synthesis. The RS domain of the PNA-peptide should provide enhancement in all splicing systems, but using it will not further understanding of how enhancers work in different disease systems. This is because the RS domain just recruits SR proteins, but in order to better understand enhancer sequences, an approach which investigates SR proteins binding to enhancer RNA sequences is required. TOES can be used to further develop the understanding of enhancer sequences and to determine whether they are pre-mRNA and/or tissue specific. It was with this in mind that the TOES system was investigated instead of the PNA-peptide system. However, for the TOES system to be suitable for enhancer sequence investigations, it needed to be adapted. By bioconjugating TOES annealer and enhancer sequences together, multiple annealer/enhancer pairs could be quickly produced without the need for resynthesizing the annealer half of the TOES to modify the enhancer half or vice versa.

3.4 – Hypothesis

Increasing the length and flexibility between an annealer and enhancer sequence will enhance splicing to exon 7 in *SMN2*, therefore increasing the levels of exon 7 inclusion.

3.5 – Results

3.5.1 – *SMN2* Splicing

3.5.1.1 – 2'OMe RNA *SMN2* Splicing

The bioconjugated tripartite sequences RNA3-RNA11 were used to test whether the presence of a bioconjugation group and varying lengths of HEG could be tolerated within a *SMN2* splicing reaction (Figure 3.3).

A splicing assay was carried out, with the reactions run simultaneously in triplicate, using NE from the same tube. Reactions were run using a master mix of 1.5mM rATP, 3.2mM MgCl₂, 20mM CrPi and 50mM KGlu and 40% NE, as before, with RNA3-11 and GGA-O at concentrations of 50nM and 100nM against a water control which contained no oligonucleotides. Reactions were run for 2 hours before being treated with proteinase K, ethanol precipitated and analysed by 6% denaturing PAGE (Figure 3.4 & Appendix Figure 5.19).

Similar levels of exon 7 inclusion were observed in most of the splicing reactions, averaging out at ~20%, with a few exceptions which shall now be discussed. At 100nM, GGA-O gave rise to 26% inclusion levels. The only difference between GGA-O and RNA3 (which gave rise to only 20% inclusion levels) and therefore the cause of this difference in inclusion levels, is the lack of a triazole group in GGA-O. RNA6 gave rise to exon 7 inclusion levels of 35% and 40% at 50nM and 100nM, respectively. RNA6 has a triazole group at the annealer end and a single HEG unit at the enhancer end. When the triazole group is moved from the annealer end to the enhancer end (which gives the structure of RNA4) exon 7 inclusion levels are reduced from 40% to 21% at 100nM and from 35% to 20% at 50nM. This reveals that the triazole group gives

rise to greater exon 7 inclusion when positioned at the annealer end. At 100nM, RNA7 gave rise to 35% exon 7 inclusion levels. In RNA7, the triazole group is positioned in the middle of the strand with a HEG unit either side, attached to the RNA strands. If the triazole group is maintained in the middle of the strand but the length of the flexible linker is increased, as in RNA11, inclusion levels are reduced back down to 20%. Increasing the flexibility from one or two units of HEG to 10 or 20 units causes a decrease in enhancer activity due to the increased entropy required for the enhancer sequence to contact the splice site and therefore to affect splice site selection. At 50nM however, RNA7 however only gave rise to exon 7 inclusion levels of 20%, the same levels as GGA-O.

The data in Figure 3.4 for RNA3-11 showed that the HEG units and the bioconjugation group were tolerated within a splicing reaction as exon 7 inclusion levels were enhanced for all tripartite sequences. This system can therefore be used for screening annealer and enhancer pairs that could potentially rectify splicing defects. The tripartite sequences also had similar or greater levels of exon 7 inclusion when compared to GGA-O indicating that the HEG and the triazole group in RNA6 and RNA7 were beneficial for enhancement of exon 7 inclusion. As both RNA6 and RNA7 had much higher activity than GGA-O and all other sequences tested, work was undertaken to further improve the efficiency of these two tripartite sequences. It was suggested that increasing protein binding to these strands, may improve exon 7 inclusion levels and therefore enable production of more effective TOES derivatives for treating SMA. An approach that has been used to increase protein binding is the incorporation of phosphothioates (pS) into 2'OMe RNA strands².

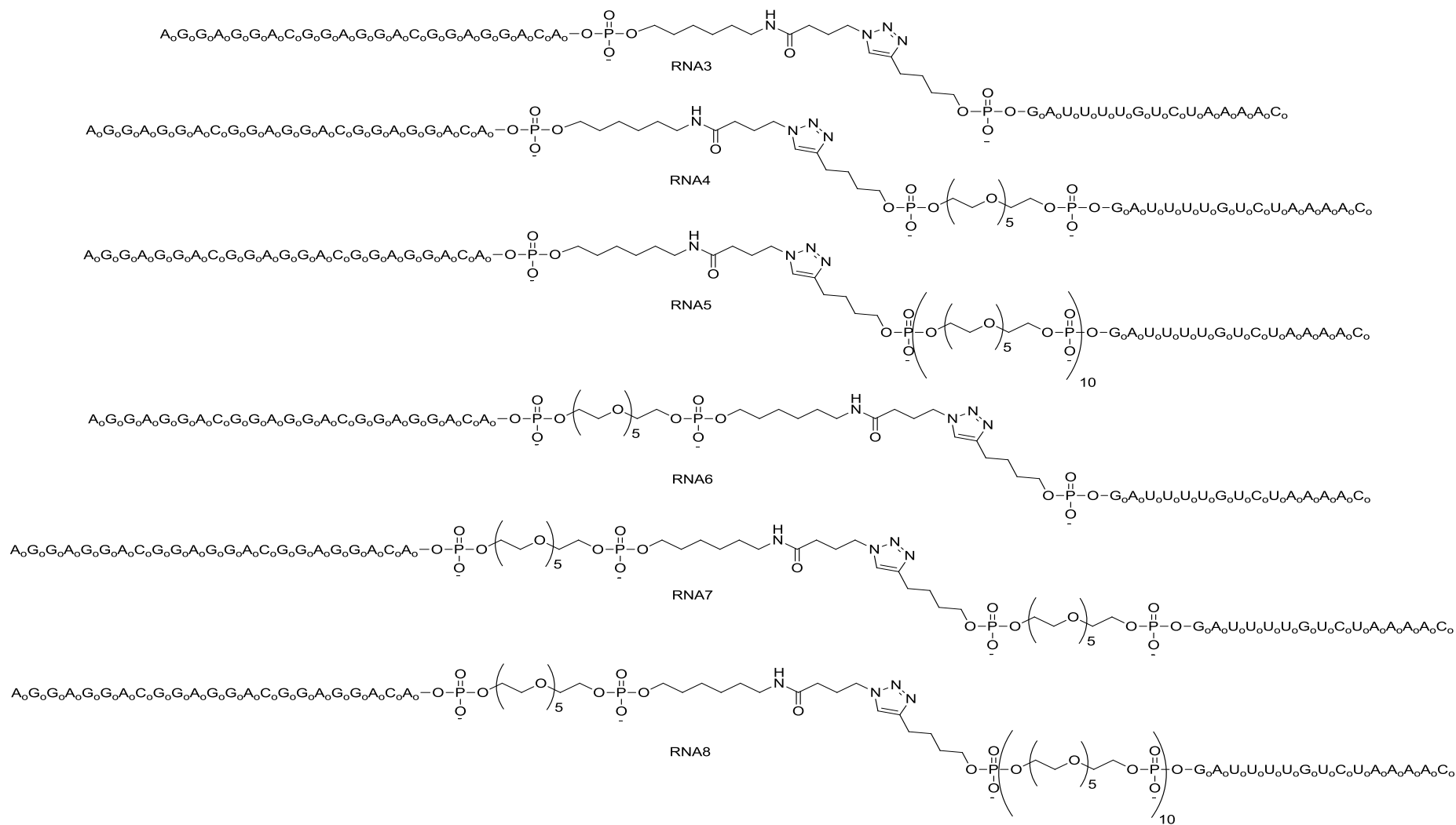


Figure 3.3. RNA3-11 tripartite sequences and GGA-O used in SMN2 splicing experiments.

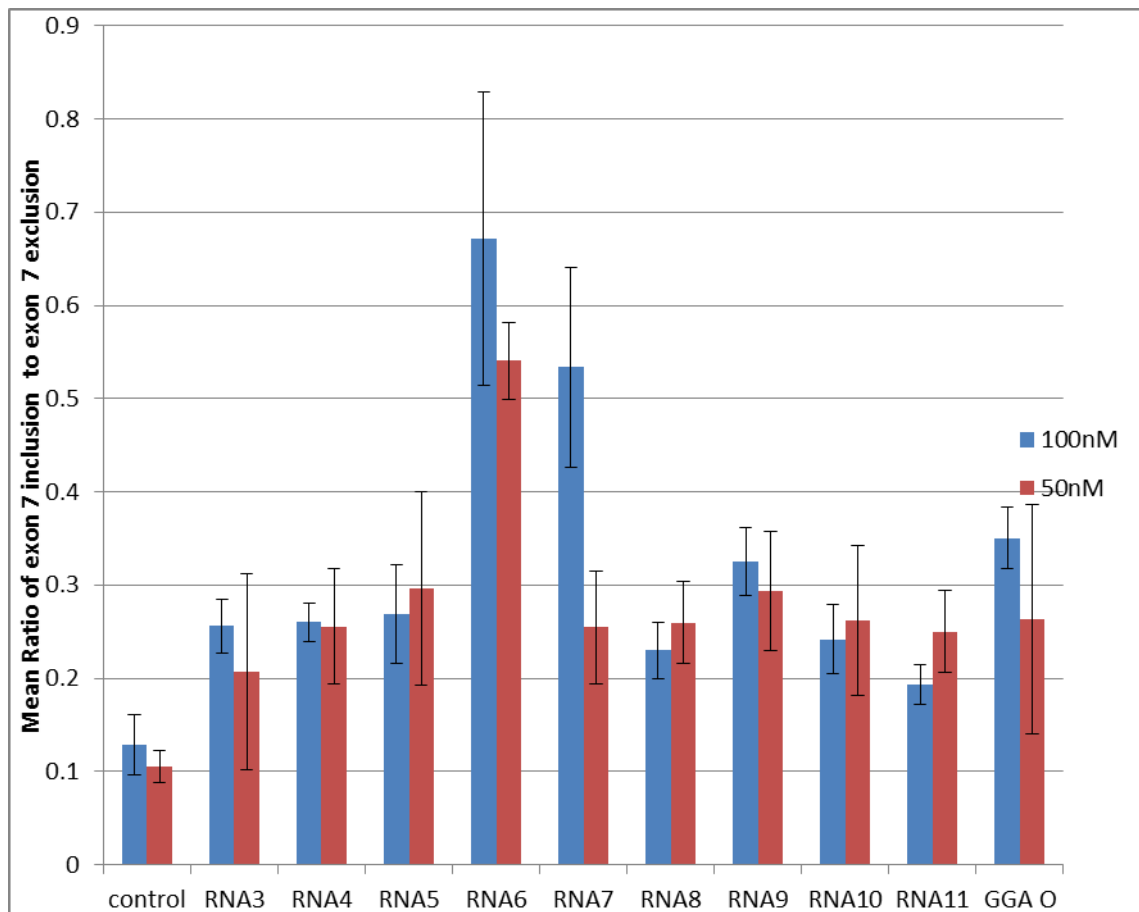


Figure 3.4. Graph displaying SMN2 mean splicing ratio of exon 7 inclusion to exon 7 exclusion when using RNA3-11 tripartite sequences and GGA-O at concentrations of 100nM and 50nM after two hours

3.5.1.2 – Increasing Protein Binding

TOES were constructed with different numbers of GGA repeats (*Section 2.1*)².

The higher the number of GGA repeats, the greater the levels of exon 7 inclusion were, indicating that the increased protein binding caused by GGA repeats is essential to enhancer activity. As the end goal is for these strands to be taken up into cells, a compromise has to be made between optimal oligonucleotide length and cellular uptake levels. By maintaining the same length strand but incorporating phosphothioate, protein binding² can be increased without negatively impacting cellular uptake. Phosphothioate versions

of RNA6 and RNA7 were synthesised to investigate whether a full pS 2'OMe backbone would enhance exon 7 inclusion (Figure 3.5).

SMN2 splicing was carried out as in *Section 4.5.1.4* with 100mM tripartite sequences run against GGA-O, GGA-S (an RNA full pS strand) and a water control. The reactions were treated with proteinase K, ethanol precipitated and analysed by 6% denaturing PAGE (Figure 3.4 & Appendix Figure 5.20).

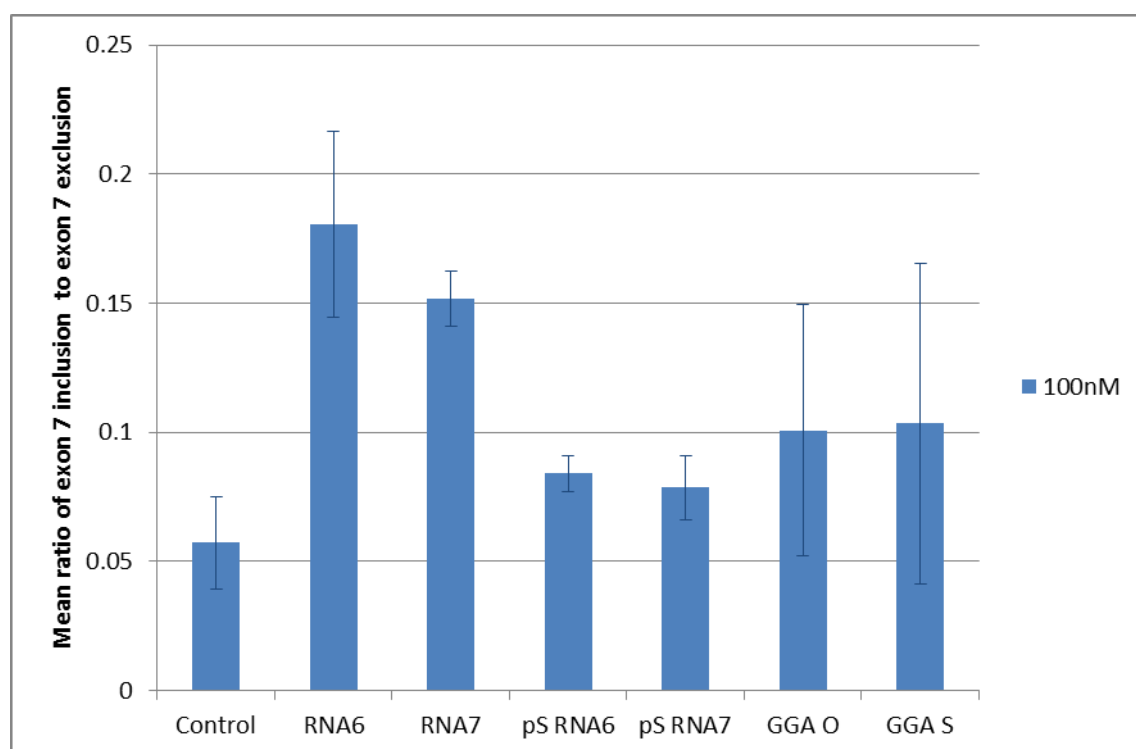


Figure 3.5. Graph displaying SMN2 mean splicing ratio of exon 7 inclusion to exon 7 exclusion when using RNA6, RNA7, pS RNA6, pS RNA7, GGA-O and GGA-S at 100nM after two hours.

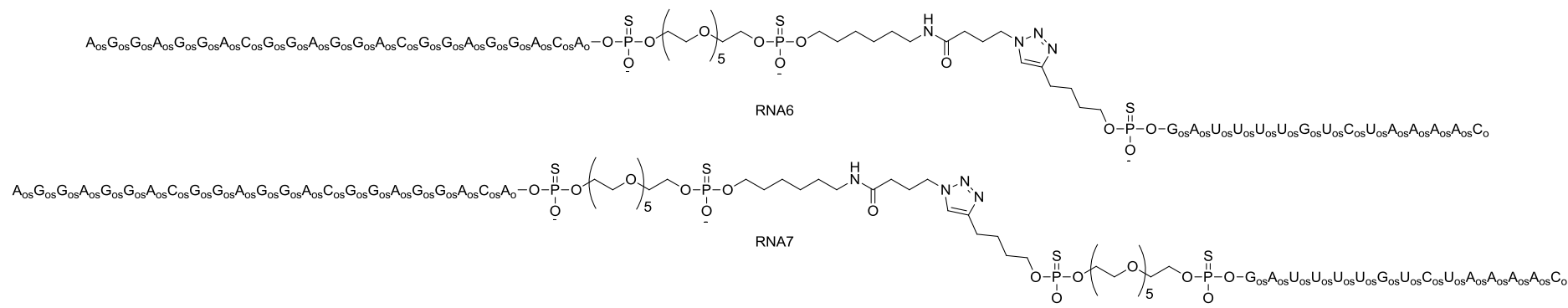


Figure 3.6. pS RNA6 and pS RNA7 tripartite sequences used in SMN2 splicing experiments.

Oligonucleotides	Mean exon 7 inclusion of spliced products (%)
	100nM oligonucleotides
Water control	5
RNA6	15
RNA7	13
pS RNA6	8
pS RNA7	7
GGA O	9
GGA S	9

Table 3.1 Mean exon 7 inclusion of spliced products (%)
in the presence of different tripartite sequences at
100nM

The data in Table 3.1 reveals that addition of a pS backbone to RNA6 decreased exon 7 inclusion levels from 15% to 8%. RNA7 and its phosphothioate counterpart, pS RNA7, followed the same pattern, with a decrease in exon 7 inclusion levels, from 13% to 7% being observed when a pS backbone was added to RNA7. RNA6 and RNA7 gave rise to higher levels of exon 7 inclusion (15% and 13% respectively) than GGA-O and GGA-S, which achieved only 9%. Overall this data reveals that a full pS backbone does not increase exon 7 inclusion levels.

To determine whether flexibility or distance between the annealer and enhancer were responsible for the high exon 7 inclusion levels observed in Figure 3.4 with RNA6, a different linker was investigated.

3.5.1.3 – Constrain Freedom of Movement of the Linker and its Impact on Splicing

To investigate the importance of the flexibility of HEG in RNA6, with regards to splice site selection, two strands, RNA14 and RNA15, were synthesised (Figure 3.7). RNA14 and RNA15 are approximately equivalent in length to RNA6, but in RNA14 both the HEG units and the triazole group are replaced with 6 abasics. In RNA15 the HEG is replaced with 3 abasics, but the triazole group is maintained. By comparing the exon 7 inclusion levels of RNA6, RNA14 and RNA15 during an *SMN2* splicing assay we can determine whether distance between the enhancer and annealer is essential for RNA6 activity or whether flexibility is what is important for splice site enhancement.

Two sets of triplicate repeats of RNA6, RNA14, RNA15, RNA2_A and GGA-O were annealed to pre-mRNA, with one set being heated at 70°C for 15 seconds and the other set at 30°C for 10 minutes, before being placed on ice. The splicing reactions were then carried out and analysed as in *Section 3.5.1.2*, using 2mM MgCl₂ and 200nM concentrations of tripartite sequences (Figure 3.8 & Appendix Figure 5.24).. 2mM MgCl₂ was used as this batch of NE had higher levels of MgCl₂ in its stock solution than previous batches of NE.

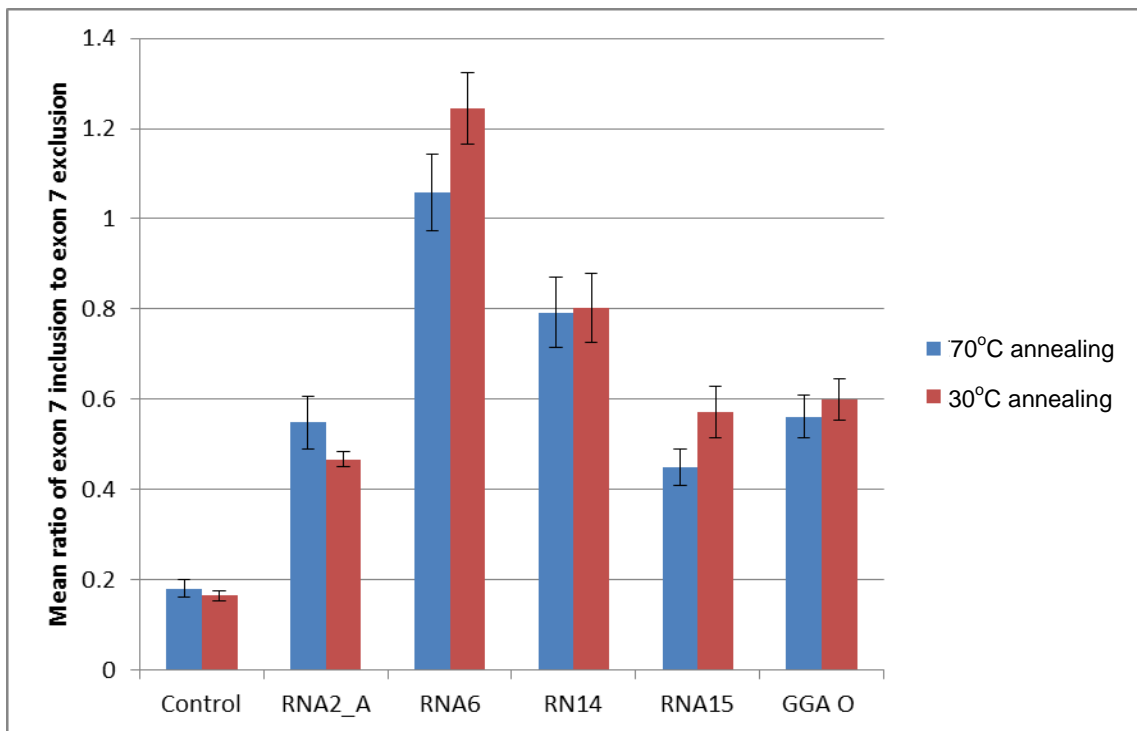


Figure 3.8. Graph displaying SMN2 mean splicing ratios of exon 7 inclusion to exon 7 exclusion when using RNA6, RNA14, RNA15, RNA2_A and GGA-O at 200nM after two hours. Annealing temperatures of 70 and 30°C were used prior to splicing.

Table 3.2. Mean exon 7 inclusion of spliced products from splicing reactions using different annealing temperatures in the presence of RNA6, RNA14, RNA15, RNA2_A and GGAO at 200nM (%)

Oligonucleotides	Mean Exon 7 inclusion of spliced products (%)	
	70°C annealing	30°C annealing
Water control	15	14
RNA2_A	35	32
RNA6	51	55
RNA14	44	44
RNA 15	31	36
GGA O	36	37

In Table 3.2 it can be seen that exon 7 inclusion levels were very similar when annealing temperatures of 70°C and 30°C were used, revealing that splicing results are barely affected by the annealing temperature. RNA6 gave rise to the highest exon 7 inclusion levels of 51-55%, whereas RNA14, RNA15 and GGA-O only achieved 44%, 36-37% and 31-36%, respectively. This data not only supports the previous data which showed RNA6 gives rise to an increase in exon 7 inclusion levels, compared to GGA-O, but it also supports the finding that RNA14 has a greater effect on exon 7 inclusion than RNA15. Therefore it can be concluded that a flexible linker is beneficial to splice site selection, as replacing the flexible linker in RNA6 with 3 abasic units, as in RNA15, reduced exon 7 inclusion to approximately the same levels as GGA-O. RNA14, which is the same length as RNA6, but contains rigid abasic linkers gave rise to high levels of exon 7 inclusion, suggesting that (contrary to earlier findings) linker flexibility is important.

3.5.1.4 – GGA Modifications and Chemistry

GGA is made up of an RNA enhancer and a 2'OMe RNA annealer with various pS placed throughout the backbone to enhance protein binding. This construct provides much greater exon 7 inclusion levels than its GGA-O counterpart² (Table 3.3). By modifying RNA6 and RNA14 constructs to include the same modifications and chemistry as GGA, these strands should, in theory, be as active, if not more active, than GGA. Unfortunately due to time restraints an RNA6 version could not be produced, but an RNA14 version was synthesised and will be referred to as RNA16. A splicing experiment was undertaken to test the activity of RNA16 against that of RNA 14. The splicing was carried out and analysed as in Section 3.5.1.6 using 2mM MgCl₂ with oligonucleotides at 50nM, 100nM and 200nM (Figure 3.9 & Appendix Figure 5.26).

Table 3.3. Mean exon 7 inclusion of spliced products for RNA14, RNA16, GGA and GGA-O (%)

Oligonucleotides	Mean exon 7 inclusion of spliced products (%)		
	200nM Oligonucleotides	100nM Oligonucleotides	50nM Oligonucleotides
control	11		
GGA	88	79	67
RNA16	86	81	64
GGA-O	30	17	10
RNA14	31	32	15

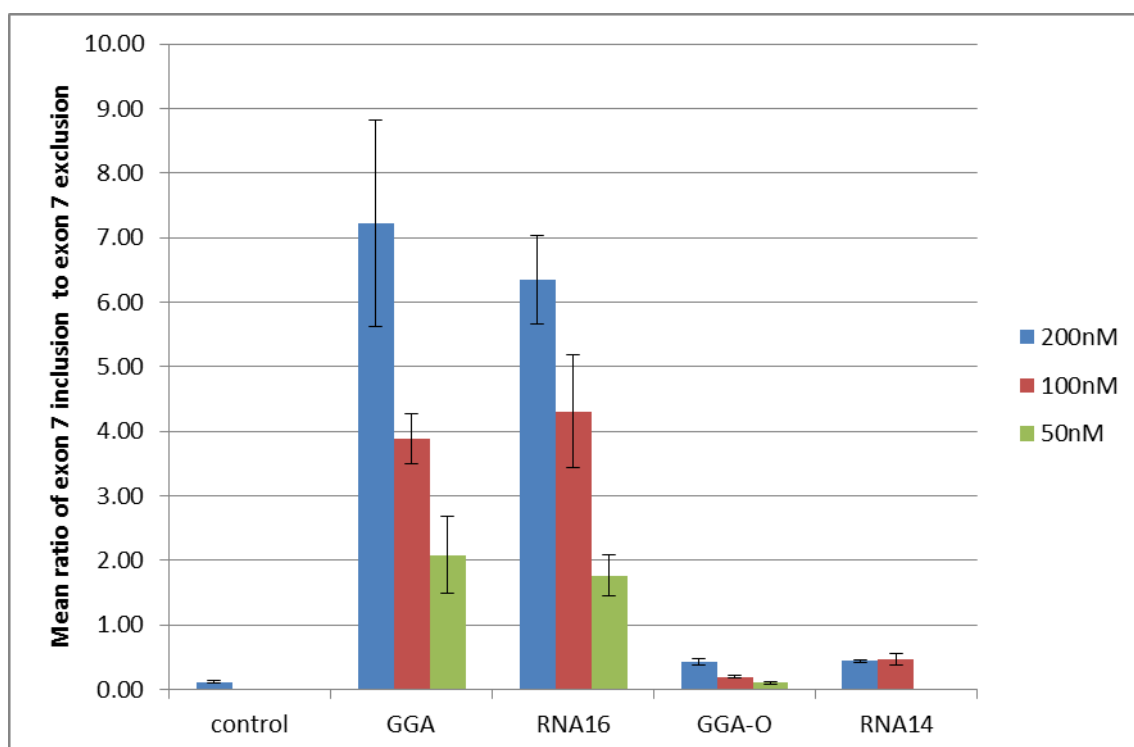


Figure 3.9. Graph displaying SMN2 mean splicing ratios of exon 7 inclusion to exon 7 exclusion when using RNA16, RNA14, GGA and GGA-O at 200nM, 100nM and 50nM after two hours.

The data in Table 3.3 shows that RNA 16 gave rise to exon 7 inclusion levels approximately equivalent to GGA, with RNA16 achieving 86%, 81% and 64% at 200, 100 and 50nM, respectively, compared to GGA which achieved 88%, 79% and 67% at 200, 100 and 50nM, respectively. When comparing RNA16 to RNA14 (2'OMe derivative), the huge effect that oligonucleotide modifications can have on splice site enhancement can be fully appreciated, as exon 7 inclusion levels rose from 31% for RNA14 to 86% for RNA16 at 200nM. Even though RNA16 gives rise to very high exon 7 inclusion levels, its increased length, resulting from the additions of the abasic linkers, means it cannot increase exon 7 inclusion levels above those achieved with GGA. It is hoped that the RNA6 version of GGA would be more active than GGA.

3.6 – Discussion

NE concentration investigations revealed that 40% NE gives rise to low levels of exon 7 inclusion, yet maintains high levels of splicing. The low levels of exon 7 inclusion, which the 40% NE caused, were ideal for enabling the effects of tripartite sequences on splice site selection to be more easily observed. Another positive reason for using NE at 40% was that less NE was required in each experiment, than the standard 50 % normally used, which significantly reduced costs when large numbers of reactions were run.

Different batches and even different tubes in the same batch of NE can cause varied levels of activity in splicing reactions¹³¹. Therefore, three replicates of each splicing reaction in each experiment were undertaken and mean values calculated, to provide a more accurate picture of the splicing activity occurring.

Exon 7 inclusion levels observed in splicing reactions with RNA3-11 were all similar to those observed with GGA-O (Table 3.1), indicating that the triazole bioconjugation group in tripartite sequences is tolerated with and without HEG units. Increasing the flexibility of the tripartite strand by addition of 1 or 2 HEG units to the tripartite sequences, as in RNA6 and RNA7, gave rise to higher exon 7 inclusion levels than were observed with GGA-O (Table 3.1). However, the addition of 10 (RNA5, RNA9), 11 (RNA8, RNA10) or 20 (RNA11) units of HEG reduced exon 7 inclusion levels. This is because the extra flexibility created by the extra HEG units reduced the probability of enhancer sequences, and therefore SR proteins, interacting with exon 7 splice sites (Table 3.1). Moving the flexible linker from the enhancer side (RNA6) to the annealer side (RNA4) diminished the increase of exon 7 inclusion (Table 3.1). A possible

reason for this could be that the triazole group, when directly next to the enhancer sequence within the tripartite sequence (RNA4) may have interfered with protein binding or reduced the looping ability of the enhancer, thereby reducing enhancer-bound SR protein interactions with exon 7 splice sites.

pS RNA6 and pS RNA7 exerted less of an effect on exon 7 inclusion than RNA6 and RNA7. As to why RNA strands which have a higher likelihood of binding to proteins actually reduce enhancement activity is unclear. It could be that the pS backbone affects the strength of the binding between pre-mRNA and annealers. If it weakens the tripartite sequence binding to the pre-mRNA, it could reduce the amount of enhancer sequence available at exon 7, therefore reducing recruitment of SR proteins to the splice site. If the pS backbone caused the binding between pre-mRNA and annealers to become too strong, tripartite sequences could possibly have blocked the splicing machinery from binding to the 3'SS. This may also have occurred during TOES optimisation studies when the strength of annealing was increased (*Section 2.1*). T_m values calculated from CD analysis for enhancer sequences, revealed that pS strands formed stronger secondary structures than non-pS strands (*Section 2.4.14*). This also suggests that pS may increase annealing strength of oligonucleotides to pre-mRNA. This all implies that the reason for the reduced activity observed with pS RNA, is quite likely to be due to binding between pre-mRNA and annealers being too strong. Further investigations would however need to be undertaken to determine exactly why these pS RNA strands were less active.

When investigating the effect that the flexibility of linkers has on splicing, a few problems occurred. The data obtained from the first experiment showed high amounts of variation, giving rise to large error bars on the graph produced. The

experiment was repeated and the results remained highly variable. Different tubes of NE had been used in both experiments, which may have accounted for the variation between reaction runs. Improper mixing of the master mix was thought to be the main cause of the variation between replicate samples within each run. The whole experiment was re-run with a new batch of NE, using NE from the same tube and the master mixes were thoroughly mixed before they were added to the splicing reactions. These changes reduced the variability of results overall, but the variability of the data obtained from RNA14 reaction samples was still high. Another cause for concern was that RNA6 and GGA-O gave rise to similar levels of exon 7 inclusion, which had not occurred in any previous experiments.

It is thought that variations in salt concentrations in the NE buffer were to blame for the unusual result of RNA6 and GGA-O giving similar exon 7 inclusion levels. Therefore, splicing experiments were carried out with the addition of 2mM and 3.2mM MgCl_2 (3.2mM being the standard MgCl_2 concentration used in splicing reactions) to investigate. Exon 7 inclusion to exon 7 exclusion ratios obtained from experiments run with the original batch of NE were more similar to the ratios obtained in the splicing experiments run with 2mM MgCl_2 than those obtained in experiments run with 3.2mM MgCl_2 . This suggests that the concentration of the MgCl_2 in the NE used in the later splicing experiments was higher than the concentration in the NE in the initial splicing reactions (*Sections 3.5.1.1 – 3.5.1.3*), which would account for the disparity between splicing results. RNA6 contains a triazole group bound to a HEG unit which provides flexibility between the enhancer and annealer sequences, whereas in RNA14 there are 6 abasic groups instead of a HEG unit, which makes it significantly

less flexible than RNA6, whilst being the same length. In Table 3.3 it can be seen that exon 7 inclusion levels of RNA6 and RNA14 were near identical (27% and 26%, respectively). This suggests that enhancement of exon 7 inclusion is more due to the extended length of the RNA strands than to the flexibility. However, comparison of RNA6 and RNA15 suggests that having a flexible linker bound to a triazole group in an RNA strand appears to increase exon 7 inclusion. This is because RNA6, which contains a HEG unit, gives rise to higher exon 7 inclusion levels than RNA15 (Table 3.3), which contains three abasic groups instead of a HEG unit and is the same length as RNA6. Therefore, it seems that both the length and flexibility of RNA strands can affect exon 7 inclusion levels in *SMN2* splicing. Variability of results were reduced by carrying out annealing reactions in the same v-wells as splicing reactions, which prevented the loss of material that would normally occur through pipetting when samples are transferred from PCR tubes. Annealing temperatures of 30°C or 70°C (*Section 3.5.1.6*) were trialled, but made very little difference to exon 7 inclusion levels and gave very similar splicing results to those obtained in *Section 3.5.1.5*. When annealing and splicing were carried out in the same v-wells of a microtitre plate, RNA6 gave rise to greater exon 7 inclusion levels than RNA14 (*Section 3.5.1.6*). However, when annealing was carried out in PCR tubes then transferred to a microtitre plate for splicing, RNA6 and RNA14 gave rise to near identical exon 7 inclusion levels (*Section 3.5.1.5*). Reactions that were annealed and spliced in the same plate wells were carried out in triplicate at two different annealing temperatures and all gave similar results (*Section 3.5.1.6*), so the results are reliable. Only one data point was obtained for reactions where annealing and splicing were carried out separately (*Section*

3.5.1.5) so the data which suggested similar exon 7 inclusion levels were achieved with both RNA6 and RNA14 can not be relied upon. It therefore seems that loss of material during sample transfer between PCR tubes and microtitre plates is the reason why RNA6 and RNA14 appeared to provide such similar levels of exon inclusion.

A splicing experiment which compared exon 7 inclusion levels between samples that were annealed in PCR tubes, then transferred to a microtitre plate against samples that were annealed and remained in a microtitre plate was conducted (*Section 3.5.1.7*). The results confirmed that loss of material during sample transfer between PCR tubes and microtitre plate v-wells negatively affects exon 7 inclusion, with the PCR-annealed samples giving rise to lower exon 7 inclusion levels (Table 3.4).

As both RNA6 and RNA14 gave higher levels of exon 7 inclusion than GGA-O, it was suspected that the GGA version of them may also increase exon 7 inclusion levels and therefore make TOES even more effective. Unfortunately due to time constraints the RNA6 version, referred to as RNA6A, was not made, but the RNA14 version, referred to as RNA16, was synthesised. Unfortunately RNA16 did not increase exon 7 inclusion to higher levels than were achieved with GGA (Table 3.6). As RNA6 had given greater exon 7 inclusion levels than RNA14, there is a good chance than RNA6A, if synthesised, would increase exon 7 inclusion to higher levels than those obtained with GGA. This work is ongoing and the outcome will be determined at a later date.

3.7 – Conclusion

Tripartite sequences are a powerful tool for screening enhancer and annealer sequences that could potentially correct for aberrant splicing in disease pathways. The addition of a bioconjugation group between two RNA strands, to form a tripartite sequence, is tolerated in splicing reactions. All 2'OMe RNA tripartite sequences tested increased exon 7 inclusion levels above basal levels. GGA tripartite sequences were even more effective at increasing exon 7 inclusion. However RNA6a, if synthesised, could potentially increase exon 7 inclusion to higher levels than those obtained with GGA, since RNA6 gave greater exon 7 inclusion levels than GGA-O. HEG units are tolerated in tripartite sequences and in the case of RNA6 and RNA7, HEG actually increased exon 7 inclusion levels. Less flexible linkers (abasic linkers) the same length as HEG gave lower levels of exon 7 inclusion, suggesting that the flexibility of the HEG linker plays an important role in the enhancement of exon 7 inclusion. Phosphothioate versions of tripartite sequences have less activity than non-phosphothioate versions. The reason for this is still under investigation, but may be due to the affect that a pS backbone has on the strength of the binding between annealer sequences and mRNA.

3.8 – Experimental

3.8.1 – Sequences

3.8.1.1 – PCR Sequence

DNA templates used for the transcription of the *SMN2* mini gene are shown in Appendix Figure 5.23.

3.8.1.2 – Pre-mRNA Sequence

The *SMN2* mini gene sequence consists of exon 2 and exon 3 of β globin with exon 7 of *SMN2* in the middle (Appendix Figure 5.24)¹³⁹. An internally truncated version of intron 2 of β globin (Appendix Figure 5.24, lower case letters in green) was cut by Sal I and exon 7 of *SMN2*, together with portions of the flanking intron sequences, was cloned into the mini-gene (Appendix Figure 5.24, highlighted in red).

3.8.2 – Standard Protocols

3.8.2.1 – ³²P-radiolabelled (Hot) Transcription Protocol

Hot transcription buffer - Tris (0.2ml, 1M, pH = 7.5), MgCl₂ (30 μ l, 1M), spermidine. HCl (11.5 μ l, 0.87M), NaCl (10 μ l, 5M) and water (0.75ml).

Low GTP rNTP mix - rATP (5 μ l, 100mM), rCTP (5 μ l, 100mM), rUTP (5 μ l, 100mM) and rGTP (0.5 μ l, 100mM).

A standard transcription reaction consisted of the following:

2 μ l Hot transcription buffer

2.5 μ l water

1 μ l Cap

1 μ l low GTP rNTP mix

1 μ l DNA template (50ng/ μ l from PCR)

0.5 μ l 0.1M DTT

These reagents were combined, vortexed to ensure thorough mixing before RNaseOUT™ (0.5 μ l) and T7 polymerase (0.5 μ l) was added. The reaction was

vortexed again before finally adding ^{32}P labelled GTP (1 μl). The reaction was mixed with a pipette and then placed in a water bath at 37°C for 1 hour.

After incubation for one hour the sample was treated with 10 μl formamide dyes and then run on a 6% denaturing polyacrylamide gel at 18W for 2 $\frac{1}{2}$ hours. The glass plates were separated and saran wrap was placed over the open side of the gel. Glow in the dark stickers placed on to the saran wrap to observe the reference points on the developed x-ray film. The gel was then exposed to an x-ray film in a dark room for 1 – 3 minutes before developing the film. The developed film was then lined up using the reference stickers and the band excised from the gel. The cut band was treated with RNA elution buffer (350 μl) and left at 5°C overnight. After incubation the sample was heated to 30°C for 10 minutes to re-dissolve the precipitated SDS before transferring the supernatant to a new tube. Ethanol (1.05ml, 3x vol of RNA elution buffer) was then added to the supernatant, vortexed and then centrifuged for 30 minutes at 13,400rpm. The supernatant was removed and the RNA pellet washed with ethanol (250 μl). The mixture was centrifuged for 15 minutes at 13,400rpm. The supernatant was once again removed and the pellet dried and re-suspended in water.

3.8.2.2 – Standard *in vitro* Splicing Reagents

A standard splicing reaction has a total volume of 10 μl and consisted of rATP, MgCl_2 , CrPi, Kglu, nuclear extract (NE), RNA and water. The final concentrations of reagents in the splicing reactions were as follows:

1.5mM rATP

3.2mM MgCl_2

20mM CrPi

50mM Kglu

These final concentrations were achieved by using the volume (per reaction) stated below of the relevant stocks:

0.3µl rATP 50mM

0.4µl MgCl₂ 80mM

0.4µl CrPi 0.5M

2µl Kglu 250mM

1µl RNA

4µl Nuclear extract (40% NE)

Make up to 10µl with water.

rATP, MgCl₂, CrPi were made up as a master mix and then 1.1µl aliquoted out into each reaction.

3.8.2.3 – Standard *in vitro* Splicing Protocol

RNA tripartite structures (2µM) were treated with ³²P labelled RNA (1µl) and Kglu (2µl, 250mM) before adding water to make up to total volume of 4.9µl. The reactions were then placed in a PCR machine and heated to 80°C followed by slow cooling over 35 minutes. After incubation the samples were placed on ice and then transferred onto 1.5ml eppendorfs. The master mix of rATP, MgCl₂ and CrPi was combined with NE (4µl, 40% NE in the final splicing mixture) and then 5.1µl to each reaction. The reactions were vortexed before being placed in a 30°C water bath for the duration of the time course. Samples taken during the time course were placed on dry ice to stop the reaction. Once the splicing reaction was complete proteinase K (10µl, 1 in 25 dilutions in PK buffer) was added to each of the reactions and heated at 37°C for 30mins. After 30 minutes

the samples were treated with ethanol (200µl) and centrifuged at 6,200rpm (microtitre plate centrifuge) for 15 minutes. The supernatants were then removed and the pellets washed with ethanol (200µl) and centrifuged again for 5mins. The supernatants were again removed and the pellets dried in a vacuum desiccator. The samples were then treated with 10µl formamide dyes before being either placed in the -80°C freezer or used directly.

The products were analysed by running a 6% denaturing polyacrylamide gel at 30W for 1³/₄ hours. The plates were then separated and the gel transferred to 3M[™] paper. The open side of the gel was covered in saran wrap before the gel was dried on a gel drier for approximately an hour. Once dried the gel was exposed to a phosphor screen overnight in a cassette. In the morning the phosphor screen was imaged using a Packard Cyclone[™] and the data quantified.

3.8.3 - PCR

dATP(10µl, 100mM), dGTP (10µl, 100mM), dTTP (10µl, 100mM) and dCTP (10µl, 100mM) were mixed together before being treated with T.E.1 (60µl). The dNTP (10mM) mix was then used in a PCR.

A standard PCR reaction was made up as outlined below.

34.2µl water

12µl 5x colourless Gotaq[™] reaction buffer

1.5µl dNTP mix (10mM)

3µl primer 1 (16) (10pmol/µl)

3µl primer 2 (17) (10pmol/µl)

0.6µl taq polymerase (5µmol/µl)

Two tubes of the above reagents were aliquoted out with one tube being treated with 2µl DNA (150ng/µl) and the other tube treated with 2µl water. The DNA added was a *SMN2* mini gene with beta globin exon 2, part of *SMN2* intron 6, *SMN2* exon 7, part of *SMN2* intron 7 and a beta globin exon 3. The reactions were then placed in a PCR machine using the PCR cycle highlighted below. The samples were heated to 95°C for 2 minutes before the PCR cycle started.

PCR cycle

1. Heat to 94°C for 15 seconds to denature
2. Cool to 50°C for 20 seconds to anneal
3. Heat to 72°C for 1½ minutes to elongate

This cycle was repeated 35 times to produce the PCR product.

A 1.5% agarose gel was run loading 1µl of PCR product in 2.5µl of dyes. As the gel showed only one spot for the PCR product and the negative control was clean the sample was worked up using a phenol/chloroform extraction.

The 50µl PCR reaction was dissolve up to 100µl with water and then treated with an equal volume of phenol: chloroform: isoamyl alcohol. The sample was then vortexed and centrifuged at 14,000rpm for 2 minutes. The aqueous layer was then transferred to a new tube, treated with phenol: chloroform: isoamyl alcohol (100µl), vortexed and centrifuged again at 14,000rpm for 2 minutes.

5x colourless Gotaq™ reaction buffer is a buffer that is made by Promega that allows the DNA after PCR reaction to be run directly on a gel without mixing with other buffers and/or loading dye. It also contains 1.5mM MgCl₂ at pH 8.5 that is required for the PCR reaction.

After the aqueous layer was transferred to a new tube for the second time the aqueous layer was treated with sodium acetate (10µl, 3M, pH 5.2) followed by the addition of ethanol (330µl, 3x vol). The sample was incubated on ice for 15 minutes before centrifuging at 13,400rpm at 4°C for 30 minutes. The supernatant was then removed and 250µl of 70% ethanol was then added and centrifuged again at 13,400rpm at 4°C for 15 minutes. The supernatant was then removed, the sample dried and then re-dissolved in 20µl water.

The concentration of the PCR product was determined to be 233.1ng/µl. 4.29µl of the PCR product was then taken and treated with 15.71µl of water to give a stock solution of 50ng/µl which could then be used in transcription reactions.

3.8.4 – *In vitro* 2'OMe RNA Splicing

3.8.4.1 – *SMN2* Transcription

Transcripts were prepared as per the described hot transcription protocol (Section 3.8.2.1).

3.8.4.2 – *SMN2* Splicing Varying the Amount of NE

Four splicing reactions were run in which the NE concentration was varied from 30% to 50%. A control was run alongside lacking NE in order to investigate the extent of RNA degradation. The reactions were set up to include the same volumes for rATP, MgCl₂, CrPi, KGlu and RNA as in the standard *in vitro* splicing protocol however the volume of NE was changed from 3µl (30% NE in the final splicing reaction) to 5µl (50% NE in the final splicing reaction) and hence the volume of water changed to maintain a 10µl splicing reaction. The reactions were placed in a 30°C water bath and 0 minutes, 15 minutes, 30 minutes, 1 hour and 2 hour time points were taken during the reaction. 2µl

samples were taken at each time point with the samples being placed on dry ice to stop the reaction. Upon completion of the time course the samples were worked up as described in the standard *in vitro* splicing protocol (Section 3.8.2.3). 40% NE showed good splicing efficiency according to 87% conversion of the pre-mRNA to the spliced product exon 7 exclusion (with no exon 7 inclusion observed). This was used for all of the remaining splicing reactions.

3.8.4.3 – SMN2 Transcription

Three *SMN2* transcripts were prepared in separate tubes as described in the hot transcription protocol (Section 3.8.2.1). The samples were heated to 37°C for 1 hour, combined followed by the addition of 30µl loading dyes. The samples were run on one lane and the product isolated as described in the hot transcription protocol (Section 3.8.2.1).

3.8.4.4 – 100nM Triplicate SMN2 Splicing

Tripartite sequences RNA3-RNA11 and controls GGA O and no oligo (water control) were aliquoted out using 0.5µl of the 2µM stocks, to give 100nM final splicing concentration. These aliquots were added to ³²P labelled RNA (1µl), KGlu (2µl, 250mM) and water (1.4µl). Annealing was carried out in a PCR machine as described in standard *in vitro* splicing protocol (Section 3.8.2.3). Only the two hour time point was taken during the splicing reaction and the splicing products were worked up as described in the standard *in vitro* splicing protocol (Section 3.8.2.3).

3.8.4.5 – SMN2 Transcription

A transcript was prepared as described in the hot transcription protocol (Section 3.8.2.1).

3.8.4.6 – 50nM Triplicate *SMN2* Splicing

Tripartite sequences RNA3-RNA11 and controls GGA O and no oligo (water control) were aliquoted out using 0.25µl of the 2µM stocks, to give 50nM final splicing concentration. These aliquots were added to before being added to ³²P labelled RNA (1µl), KGlu (2µl, 250mM) and water (1.65µl). Annealing was carried out in a PCR machine as described in standard *in vitro* splicing protocol (Section 3.8.2.3). Only the two hour time point was taken during the splicing reaction and the splicing products were worked up as described in the standard *in vitro* splicing protocol (Section 3.8.2.3).

3.8.5 – *In vitro* 2'OMe pS RNA Splicing

3.8.5.1 – *SMN2* Transcription

A transcription reaction was prepared which had six reactions in one, giving transcription volumes of:

12µl Hot transcription buffer

15µl water

6µl Cap

6µl low GTP rNTP mix

6µl *SMN2* mini gene (50ng/µl from PCR)

3µl 0.1M DTT

3µl RNaseOUT™

3µl T7 polymerase

6µl ³²P GTP

The reaction was carried out as described in the hot transcription protocol (Section 3.8.2.1).

3.8.5.2 – pS Splicing

Tripartite sequences RNA6, RNA7, pS RNA3 and pS RNA4 and controls GGA, GGA S and GGA O were aliquoted out using 0.5µl of the 2µM stocks, to give 100nM final splicing concentration. These aliquots were added to ³²P labelled RNA (1µl), KGlu (2µl, 250mM) and water (1.4µl). The annealing was carried out in a PCR machine as described in standard *in vitro* splicing protocol (Section 3.8.2.3). Only the two hour time point was taken during the splicing reaction and the splicing products were worked up as described in the standard *in vitro* splicing protocol (Section 3.8.2.3).

3.8.6 – *In vitro* 2'OMe Abasic Splicing

3.8.6.1 – *SMN2* Transcription

A transcription reaction was prepared which had three reactions in one, giving transcription volumes of:

6µl Hot transcription buffer

7.5µl water

3µl Cap

3µl low GTP rNTP mix

3µl *SMN2* mini gene (50ng/µl from PCR)

1.5µl 0.1M DTT

1.5µl RNaseOUT™

1.5µl T7 polymerase

3µl ³²P GTP

The reaction was carried out as described in the hot transcription protocol (Section 3.8.2.1).

3.8.6.2 – 70°C Annealing

Tripartite sequences RNA6, RNA7, RNA14 and RNA15 and controls GGA O and no oligo (water control) were aliquoted out using 0.5µl of the 2µM stocks, to give 200nM final splicing concentration. These aliquots were then added to ³²P labelled RNA (0.5µl) and annealed in a v-well microtitre plate. The plate was heated to 70°C for 15 seconds before adding KGlu (1µl, 250mM) and water (0.45µl) and leaving to cool for 15 minutes on ice. A master mix of rATP (6.3 µl, 50mM), MgCl₂ (5.2 µl, 80mM), CrPi (8.4µl, 0.5M) and water (3.2µl) was made and 0.55µl of the master mix was aliquoted into each of the reactions. The reactions were vortexed before being treated with NE (2µl), vortexed again and then placed in a 30°C oven. Only the two hour time point was taken during the splicing reaction and the splicing products were worked up as described in the standard *in vitro* splicing protocol (Section 3.8.2.3)

3.8.6.3 – 30°C Annealing

Tripartite sequences RNA6, RNA7, RNA14 and RNA15 and controls GGA O and no oligo (water control) were aliquoted out using 0.5µl of the 2µM stocks, to give 200nM final splicing concentration. These aliquots were then added to ³²P labelled RNA (0.5µl) and KGlu (1µl, 250mM). The samples were annealed by heating to 30°C in a v-well microtitre plate for 10 minutes before cooling on ice. A master mix of rATP (6.3 µl, 50mM), MgCl₂ (5.2 µl, 80mM), CrPi (8.4µl, 0.5M) and water (22.1µl) was made and 1µl of the master mix was aliquoted into each of the reactions. The reactions were vortexed before being treated with NE (2µl), vortexed again and then placed in a 30°C oven. Only the two hour time point was taken during the splicing reaction and the splicing products were

worked up as described in the standard *in vitro* splicing protocol (Section 3.8.2.3).

3.8.7 – *In vitro* RNA Splicing

3.8.7.1 – *SMN2* Transcription

A transcription reaction was prepared which had three reactions in one, giving transcription volumes of:

6µl Hot transcription buffer

7.5µl water

3µl Cap

3µl low GTP rNTP mix

3µl *SMN2* mini gene (50ng/µl from PCR)

1.5µl 0.1M DTT

1.5µl RNaseOUT™

1.5µl T7 polymerase

3µl ³²P GTP

The reaction was carried out as described in the hot transcription protocol (Section 3.8.2.1).

3.8.7.2 – RNA16 Splicing

Tripartite sequences RNA14 and RNA16 and controls GGA *, GGA O and no oligo (water control) were aliquoted out using 0.13µl, 0.25µl and 0.5µl of the 2µM stocks to give 50nM, 100nM and 200nM final splicing concentrations respectively. These aliquots were then added to before being added to ³²P labelled RNA (0.5µl) and water (0.9µl (50nM reaction), 0.78µl (100nM reaction)

and 0.53µl (200nM reaction)) and heated in an oven at 30°C for 10 minutes. A master mix of rATP (6.3 µl, 50mM), MgCl₂ (5.2 µl, 80mM), CrPi (8.4µl, 0.5M) and KGlu (42µl, 250mM) was made and 1.47µl of the master mix was aliquoted into each of the reactions. The reactions were vortexed before being treated with NE (2µl), vortexed again and then placed in a 30°C oven. Only the two hour time point was taken during the splicing reaction and the splicing products were worked up as described in the standard *in vitro* splicing protocol (Section 3.8.2.3).

Chapter 4 – Rectifying splicing defeats using polyconjugated GNP

4 – Introduction

Alternative splicing is a system which requires the combination of U specific proteins, U RNA and other non-specific proteins, in order to determine real SSs over cryptic SSs (*Section 1.1*). It has been estimated that tissue specific alternative splicing requires between 12 and 19 different cues, to ensure correct SS selection¹⁷². Artificial ESEs in TOES and tripartite sequences require a combination of factors to stimulate splicing at a weak 5' SS. For example, RNA enhancer sequences are required to recruit SRSF-1 protein which in turn recruits U1 snRNP to the 5' SS (*Section 1.1.5.1*). The combination of RNA/RNA, RNA/protein and protein/protein interactions is essential for many processes, especially splicing.

TOES are constructed from an annealer and an enhancer sequence. The annealer sequence binds to the +2 to +16 of exon 7 in pre-mRNA^{2,139}. The enhancer sequence contains six repeats of a GGA motif for binding SRSF-1. These constructs are thought to enhance SS recognition via a looping model in which the SRSF-1 bound to the enhancer sequence on the TOES directly interact with U1 snRNP at the 5'SS (Figure 4.1)².

A monovalent system consists of a single molecule which binds at a particular site. A multivalent system consists of multiple molecules bound to a surface which then binds to multiple sites. In the monovalent TOES system, exon 7 inclusion levels can be increased by adding more GGA repeats to TOES constructs (*Section 2.1.2*) and also by increasing the concentration of TOES. However, increasing the number GGA repeats would increase enhancer length, which would negatively affect cellular uptake. Also in this monovalent system,

increasing the concentration of tripartite sequences in *SMN2* splicing reactions can decrease overall splicing efficiency (*Section 3.5.1.2*). We hypothesized that a multi-valent system could overcome these problems. Attaching multiple enhancer and/or annealer strands to a surface to create a multivalent system, would increase the effective concentration of the strands at a particular site without increasing their overall concentration in an attempt to prevent splicing inhibition. A multivalent system will also increase the number of enhancer sequences which will indirectly increase the number of GGA repeats at a particular site, thereby enabling more SRSF-1 proteins to bind at that particular site ². Using a multivalent system in which both the effective concentration and number of SRSF-1 binding sites are increased could potentially increase exon 7 inclusion levels in *SMN2*. Gold Nanoparticles (GNPs) are an ideal molecule to be used as the binding surface in a multivalent system as they can be easily attached to annealers and enhancers via Au-S interactions, forming polyconjugated GNPs and because they cause no detectable toxicity within human cells or mouse models ¹⁷³. Other benefits of using GNPs are that they can be used as vehicles for cellular uptake and they are nuclease resistant ¹⁷³. Multivalent systems comprising DNA or RNA conjugated GNPs have been used to silence genes. This has been done by either delivering small interfering (si) RNA strands into cells to promote mRNA degradation by endonucleases and exonuclease, or by using GNPs conjugated to DNA, RNA or LNA strands which bind mRNA and prevent it from being translated into protein ^{173–179}. GNP-DNA/RNA conjugates have also been used to detect specific RNA sequences and biomarkers by forming aggregates or releasing fluorescently labelled DNA strands (which are quenched while bound to GNPs) ^{10–16}.

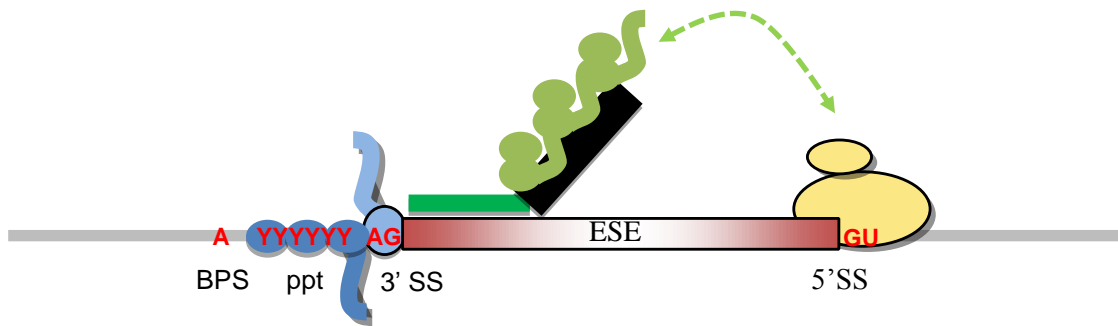


Figure 4.1. Looping model for tripartite sequences stimulating U1 recruitment. Grey line represents an intron, red box represents an exon, dark green and black lines represent the annealer and enhancer of TOES, respectively. Green circles with tails depict SRSF-1, green dotted line depicts SRSF-1 interacting with U1 snRNP, cream circles represent U1 snRNP and light and dark blue shapes represent U2AF35 and U2AF65 respectively. Red letters indicate key bases for splicing.

All approaches using multivalent GNP DNA/RNA systems to affect gene expression only interfere with the mRNA produced after splicing has occurred. At this stage no research has been carried out to investigate the effects of using polyconjugated GNP in splicing reactions and to discover whether annealer and enhancer conjugated GNP can successfully modify splicing events by either inhibiting or enhancing the inclusion of exons to the spliced mRNA.

4.1 – Aims

Optimisation studies carried out by Ian Eperons lab² revealed that increasing the number of GGA repeats and increasing the concentration of TOES improved levels of exon 7 inclusion within *SMN2*. The downside of increasing the number of GGA repeats is that it will make them more negatively charged and therefore decrease cellular uptake of TOES. As TOES constructs are

thought to act via a looping mechanism, increasing the number of strands at a particular point, by binding enhancer strands to a surface, will have a similar effect to increasing the number of GGA repeats, resulting in an increase in the number of binding sites for SRSF-1. The aim of this nanoparticle work is to investigate the effect of binding enhancer and annealer sequences to GNPs in *SMN2* splicing reactions and whether this multivalent system would be more efficacious than a monovalent (TOES) system. Further aims are to investigate whether GNP size will affect splicing and whether the presence of an annealing group on the GNPs will increase exon 7 inclusion levels.

4.2 – Hypothesis

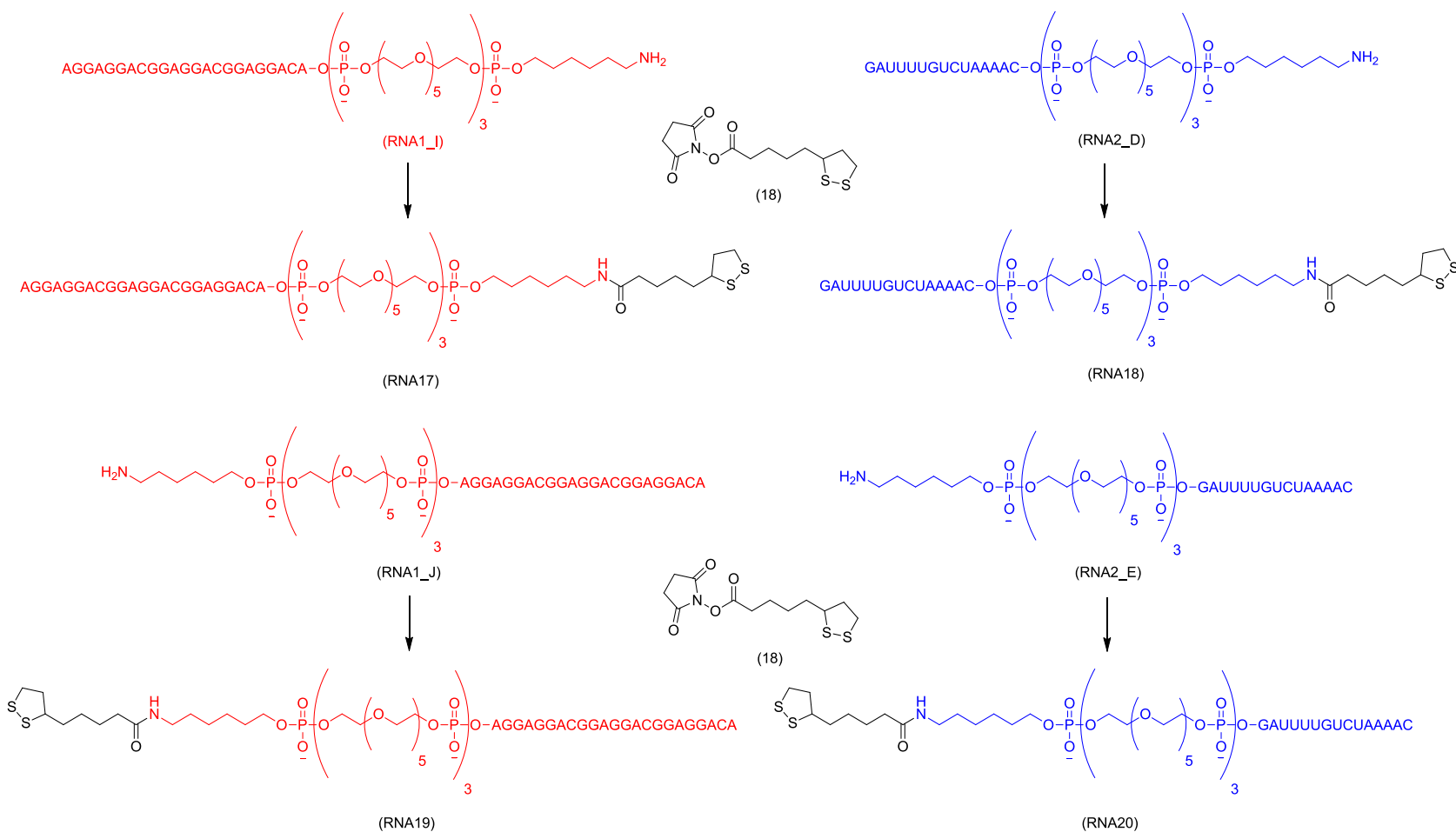
Increasing the effective concentration of enhancer strands, via conjugation to GNPs, will result in an increase in exon 7 inclusion levels in splicing reactions. Also, increasing GNP size will increase inhibition of splicing due to the steric interference of GNPs preventing splicing machinery from binding to key sites in pre-mRNA.

4.3 – Strategy

GNP will be conjugated with the same enhancer and annealer sequences that were used to produce tripartite sequences in *Section 2.2*, but with the addition of three HEG units between the 2'OMe sequences and the amino modifier. HEG units are to be used since in Chapter 3 they were shown to increase the effect of the enhancer strand on inclusion of exon 7. Another reason for using HEG units is to maximise DNA loading on GNPs (which has been previously done by Chad Mirkins group¹⁸⁷) since HEG units will reduce the amount of charge at the

end of the strands that are attached to a GNP. This is because three HEG units are approximately the same length as ten nucleotides, but contain less phosphodiester groups and therefore possess a lower charge. This means there will be less electrostatic repulsion between strands attached to a GNP allowing more strands to bind to it. Therefore the HEG units will cause enhancer and annealer sequences to be spaced away from GNPs, reducing steric crowding around the GNPs, allowing the strands to bind proteins and/or RNA. The HEG units will increase the number of strands that can be polyconjugated to GNPs (in DNA-GNP based systems) 3.1 fold compared to nucleotide spacers of approximately the same length ¹⁸⁷. As maximising RNA-GNP polyconjugation has not been investigated before, the addition of three HEG units to annealer and enhancer sequences is a good starting point.

Enhancer and annealer sequences will be prepared by solid phase synthesis and amino modifiers attached to them. As polyconjugated GNPs have not been used in splicing reactions before, the directionality of the strand required for activity is unknown. Two annealer sequences, RNA2_D and RNA2_E, and two enhancer sequences, RNA1_I and RNA1_J, will be made, with RNA2_E and RNA1_J being amino and HEG modified at the 5' end and RNA2_D and RNA1_I being amino and HEG modified at the 3' end (Scheme 4.1). These sequences will then undergo an NHS coupling reaction with compound **18** to produce RNA17-20 sequences that are ready for polyconjugation (Scheme 4.1). Compound **18** contains a dithiol which will be used to attach the 2'OMe RNA strands to the GNPs, forming a stronger attachment than would be achieved with a single thiol.



Scheme 4.1. Reactions between amino modified RNA1_I, RNA2_D, RNA1_J, RNA2_E and compound **18**, producing lipoic acid versions of the amino modified strands written 5' to 3'. Enhancers are red and annealers are blue.

GNP sizes of 5nm, 10nm, 18nm and 20nm were chosen for these investigations. Particles larger than 20nm were not chosen as it was thought they would inhibit splicing by sterically hindering protein binding to key sites. GNPs smaller than 5nm were not investigated since they would undergo rapid renal clearance in the body ¹⁸⁸. 18nm particles were prepared using a literature protocol¹⁸⁹ and 20nm, 10nm and 5nm particles were purchased from BBI International. RNA17-20 were treated with TCEP prior to polyconjugation to GNP (Figure 4.2).

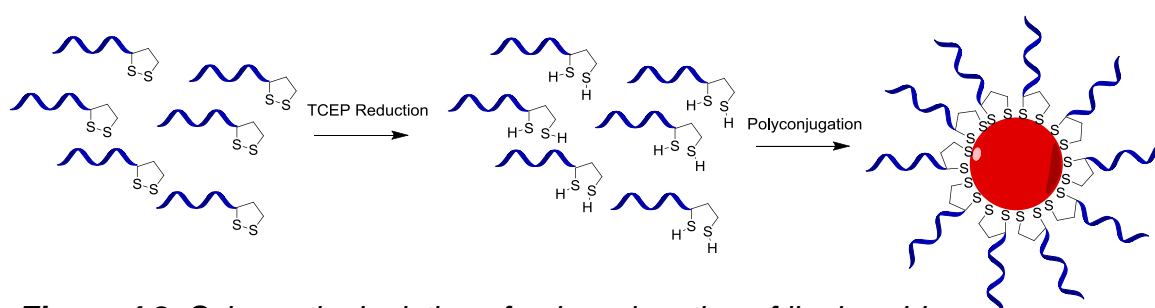


Figure 4.2. Schematic depiction of polyconjugation of lipoic acid functionalised 2'OMe RNA ON 17-20 to GNP. Blue strands represent RNA17-20.

The effect of polyconjugated GNP on *SMN2* splicing will be investigated first using 18nm GNPs. The polyconjugation coating that will be used on 20nm, 10nm and 5nm GNPs will be decided upon based on the findings from the 18nm GNP investigations.

4.4 – Results

4.4.1 – Preparation of Strands for Polyconjugation

RNA1_I, RNA1_J, RNA2_D and RNA2_E were prepared using standard 2'OMe RNA solid phase synthesis with synthesis yields of 42%, 54%, 59% and 56%,

respectively (Figure 4.3). RNA1_I, RNA1_J, RNA2_D and RNA2_E were confirmed by accurate mass (Table 4.1).

Table 4.1. *Molecular weights of RNA1_I, RNA1_J, RNA2_D and RNA2_E synthesised on an ABI 394 synthesiser. Accurate mass data included in appendix (Figure 5.27)*

Sequence Name	Expected MW	Observed MW [M+H]
RNA1_I	9163.62	9163.60
RNA1_J	9163.62	9163.49
RNA2_D	6143.60	6144.14
RNA2_E	6143.60	6143.09

4.4.2 – Conjugation of Lipoic Acid

Conjugation with lipoic acid was achieved by the addition of compound (**18**) to an aqueous solution of RNA1_I, RNA1_J, RNA2_D and RNA2_E (pH 8) (Scheme 4.1). Incubation overnight afforded compounds RNA17, RNA18, RNA19 and RNA20 (Figure 4.4). RNA17-20 were worked up using an Illustra G25 column to remove unreacted compound **18** before being characterised by mass spec (Table 4.2) and used crude in polyconjugation reactions. Not all of the starting materials were converted (Appendix Figure 5.28) however, since only the lipoic acid-conjugated strands would polyconjugate to the GNPs and the strands would be required in excess, the reactions were continued with these impurities

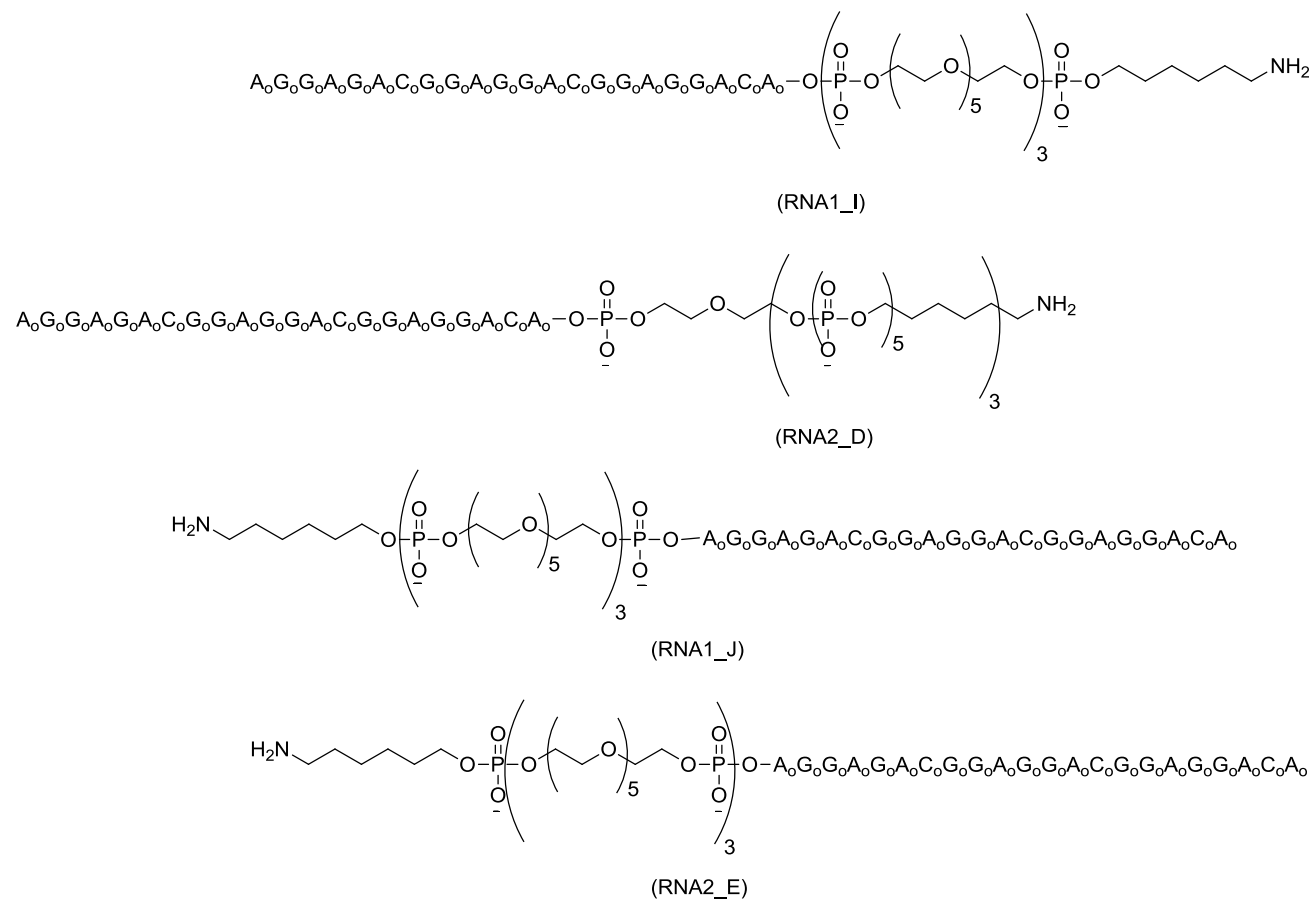
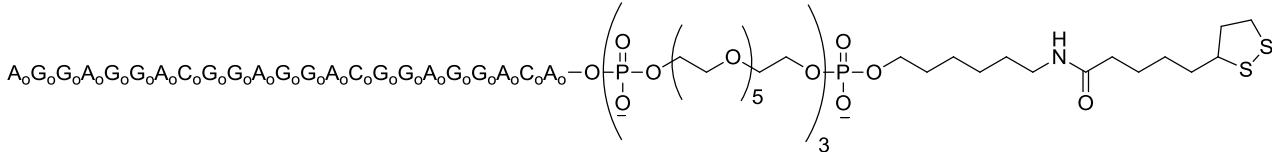


Figure 4.3. 2'OMe RNA ONs prepared by solid phase synthesis, written 5' to 3'.



192

Table 4.2. *Molecular weights of RNA17, RNA18, RNA19 and RNA20 and their starting materials. Accurate mass data included in appendix (Figure 5.28)*

Sequence Name	Expected MW	Observed MW [M+H]	Starting Material Expected MW	Observed starting material MW [M+H]
RNA17	9351.63	9352.73	9163.62	9162.77
RNA18	6331.61	6331.17	6143.60	6143.15
RNA19	9351.63	9350.90	9163.62	9163.84
RNA20	6331.61	6331.11	6143.60	6196.13

The reaction forming RNA20, gave rise to two peaks when analysed by HPLC. The first peak had a similar retention time to that of RNA2_E (Appendix Figures 5.27) however, the molecular weight (MW) observed was greater than expected (Table 4.2). As RNA2_E was found to be the correct mass when it was synthesised (Table 4.1) and because the product, RNA20 (which gave rise to the second HPLC peak) was the correct mass, the reaction was continued.

4.4.3 – Preparation of GNPs for Polyconjugation

20nm, 10nm and 5nm GNPs were purchased from BBI International. 18nm GNPs were prepared by Dr Vanessa Bonnard according to a published procedure using citrate coated GNPs¹⁸⁹. Citrate coated GNPs were first ligand exchanged with BSPP followed by ligand exchange with lipoic acid conjugated strands RNA17-20 (Figure 4.4)¹⁹⁰. GNPs were ligand exchanged with BSPP, to increase their stability in NaCl, which allowed the salt ageing step of the polyconjugation reaction to be reduced from 40 hours to 12 hours as salt additions could then be made more frequently (every 20 minutes)^{187,190}. Salt

ageing neutralises the electrostatic repulsion between RNA strands allowing a greater coating of strands to be formed on the GNP surface.

In situ reduction of RNA17, RNA18, RNA19 and RNA20 disulphide bonds using TCEP, was carried out prior to polyconjugation of the strands to GNPs (Table 4.3, GNP1-18). Polyconjugates (GNP 1-18) were obtained by the addition of 10-30µl of 2M NaCl in 0.5x TBE, 0.015% SDS up to 700mM and incubated overnight ¹⁸⁷. GNPs1-18 were centrifuged five times, with supernatants being removed after each spin, before resuspending the GNPs, to remove any unconjugated RNA, affording GNPs1-18 which were used in splicing (Figure 4.5)

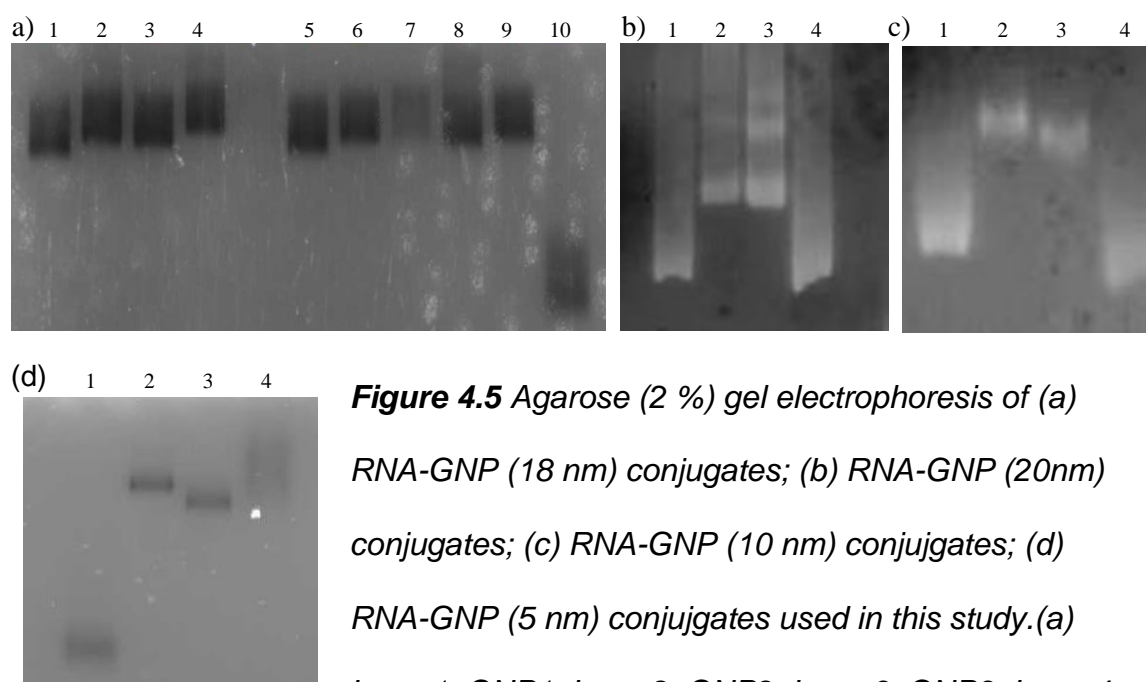


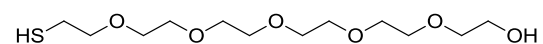
Figure 4.5 Agarose (2 %) gel electrophoresis of (a) RNA-GNP (18 nm) conjugates; (b) RNA-GNP (20nm) conjugates; (c) RNA-GNP (10 nm) conjugates; (d) RNA-GNP (5 nm) conjugates used in this study. (a) Lane 1: GNP1. Lane 2: GNP2. Lane 3: GNP3. Lane 4: GNP4. Lane 5: GNP5. Lane 6: GNP6. Lane 7: GNP6. Lane 8: GNP7. Lane 9: GNP8. Lane 10: GNP9. (b) Lane 1: BSPP coated 20nm GNPs. Lane 2: GNP10. Lane 3: GNP11. Lane 4:GNP12. (c) Lane 1: BSPP coated 10nm GNPs. Lane 2: GNP13. Lane 3: GNP14. Lane 4:GNP15. (d) Lane 1: BSPP coated 5nm GNPs. Lane 2: GNP16. Lane 3: GNP17. Lane 4:GNP18.

GNP4. Lane 5: GNP5. Lane 6: GNP6. Lane 7: GNP6. Lane 8: GNP7. Lane 9: GNP8. Lane 10: GNP9. (b) Lane 1: BSPP coated 20nm GNPs. Lane 2: GNP10. Lane 3: GNP11. Lane 4:GNP12. (c) Lane 1: BSPP coated 10nm GNPs. Lane 2: GNP13. Lane 3: GNP14. Lane 4:GNP15. (d) Lane 1: BSPP coated 5nm GNPs. Lane 2: GNP16. Lane 3: GNP17. Lane 4:GNP18.

Table 4.3. Particle size and strand coating of GNP1-18

GNP Name	GNP1	GNP2	GNP3	GNP4	GNP5	GNP6	GNP7	GNP8	GNP9	GNP10	GNP11	GNP12	GNP13	GNP14	GNP15	GNP16	GNP17	GNP18
Particle size (nm)	18	18	18	18	18	18	18	18	18	20	20	20	10	10	10	5	5	5
Strands attached	RNA 20	RNA 19	RNA 18	RNA 17	RNA 19 & RNA 20	RNA 17 & RNA 20	RNA 18 & RNA 19	RNA 17 & RNA 19	Peg-SH	RNA 17	RNA 17 & RNA 20	Peg-SH	RNA 17	RNA 17 & RNA 20	Peg-SH	RNA 17	RNA 17 & RNA 20	Peg-SH

PEG-SH =



4.4.4 – 5nm GNP16-18 Characterisation

4.4.4.1 – CPS Characterisation of GNP16-18

Characterisation of polyconjugated GNPs were only carried out on exemplar candidates GNP16, GNP17 and GNP18. GNP16 and GNP17 have 2'OMe RNA coatings, with GNP16 having an RNA17 coating and GNP17 having a mixed monolayer RNA17 and RNA20 coating. GNP18 is a PEGylated control particle which was used as a reference for the polyconjugated 2'OMe GNPs. Successful attachment of 2'OMe RNA ligands onto the surface of GNPs (GNP16 and GNP17) were confirmed by Dr Željka Krpetić using differential centrifugal sedimentation (CPS) ^{191,192}. Shifts of 1.9nm and 1.6nm were detected for GNP16 and GNP17, respectively compared to GNP18 ($d_{CPS} = 5.4\text{nm}$) as a result of overall particle density change due to the RNA attachment (Figure 4.6). Moreover, no aggregation of the particles was detected in the size range examined (2-50 nm) confirming stable colloidal dispersions of RNA functionalised GNPs (Figure 4.6).

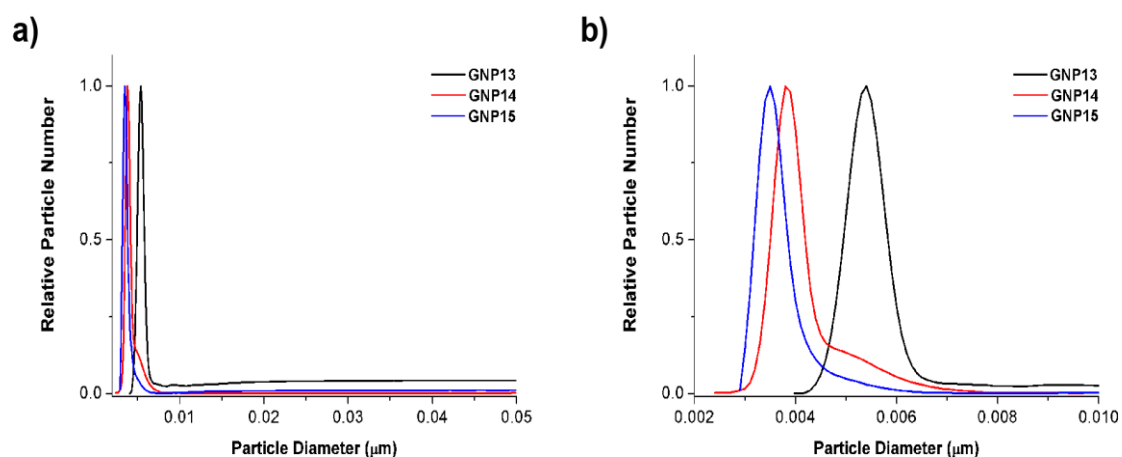


Figure 4.6. Differential centrifugal sedimentation (CPS) particle size measurements showing distributions of gold nanoparticles for GNP16 GNP17 (red) and GNP18 (black) measured in **a)** 2-50 nm and **b)** 2-10 nm size range.

4.4.4.2 – Determining Loading of GNP16-18

ICP-AES analysis was carried out by Dr Željka Krpetić to determine the ratio of Au to S. This ratio provided the total number of RNA ligands on the surface of each GNP¹⁹¹. This process involves acid digest of the GNP conjugates in *aqua regia* for 72 hours to break molecules down to single atoms. The ICP-AES ionises these atoms, creating ions of Au and S that are in excited states. The amount of material present is determined by the light energy emitted when the ions return to the ground state. This light energy is quantified by calibrating ICP-AES with a known concentration (ppm) of Au and S, providing ppm concentrations of the Au and S ions present. ICP-AES can calculate the concentrations of ions in samples containing more than one element as different elements emit light with different wavelengths, due to differences between their excited state and ground state. This method was used to calculate that 400-420 RNA ligands coated the surface of GNP16, 470-490 RNA ligands coated the surface of GNP17 and 880-900 PEG-thiol ligands coated the surface of GNP18. GNP18 loading was found to be significantly higher than that of GNP16 and GNP17. This is thought to be due to GNP18 having a PEG-thiol coating which is not charged, therefore no electrostatic repulsion has to be overcome when loading the strands onto the GNP. However, GNP16 and GNP 17 have an RNA coating which is charged and therefore electrostatic repulsion occurs which reduces the amount of loading that is possible. Another possible reason is that RNA strands are conjugated to GNPs using a di-thiol to form GNP16 and GNP17, whereas the PEG strands only conjugate to the GNP by a single thiol to form GNP18, so there is more space for more strands to bind.

4.4.4.3 – TEM Analysis of GNP16-18

TEM analysis carried out by Dr Željka Krpetić confirmed narrow size distribution of GNP16-18 (Figure 4.7). The core size of GNP16-18 were determined by counting populations of TEM samples with approximately 150 GNPs, which provides a sample that is indicative of the whole population. Results from the TEM revealed that GNPs, when loaded, increased in size from ~ 5nm to 5.77nm for GNP16, 5.57nm for GNP17 and 5.65nm for GNP18 (Figure 4.7).

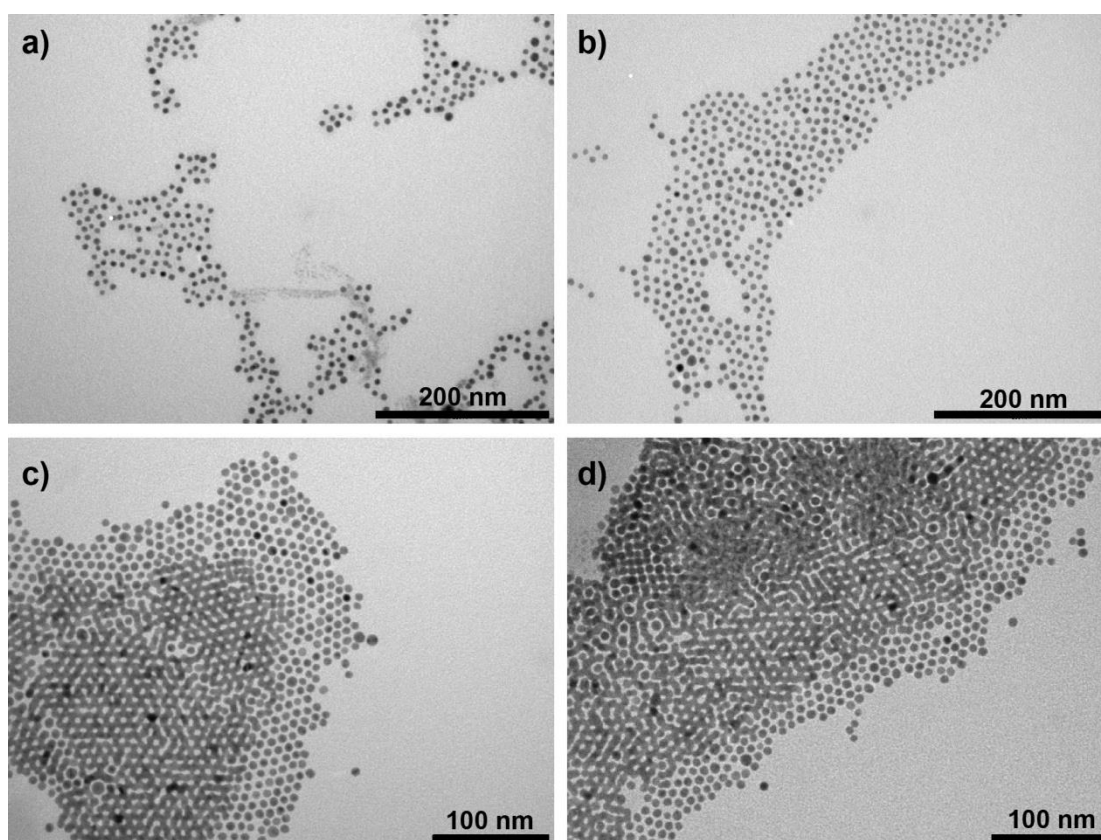


Figure 4.7. TEM micrographs of **a)** PEGylated GNPs (GNP18) and RNA-functionalized GNPs **b)** GNP17 **c)** GNP16 **d)** GNP16.

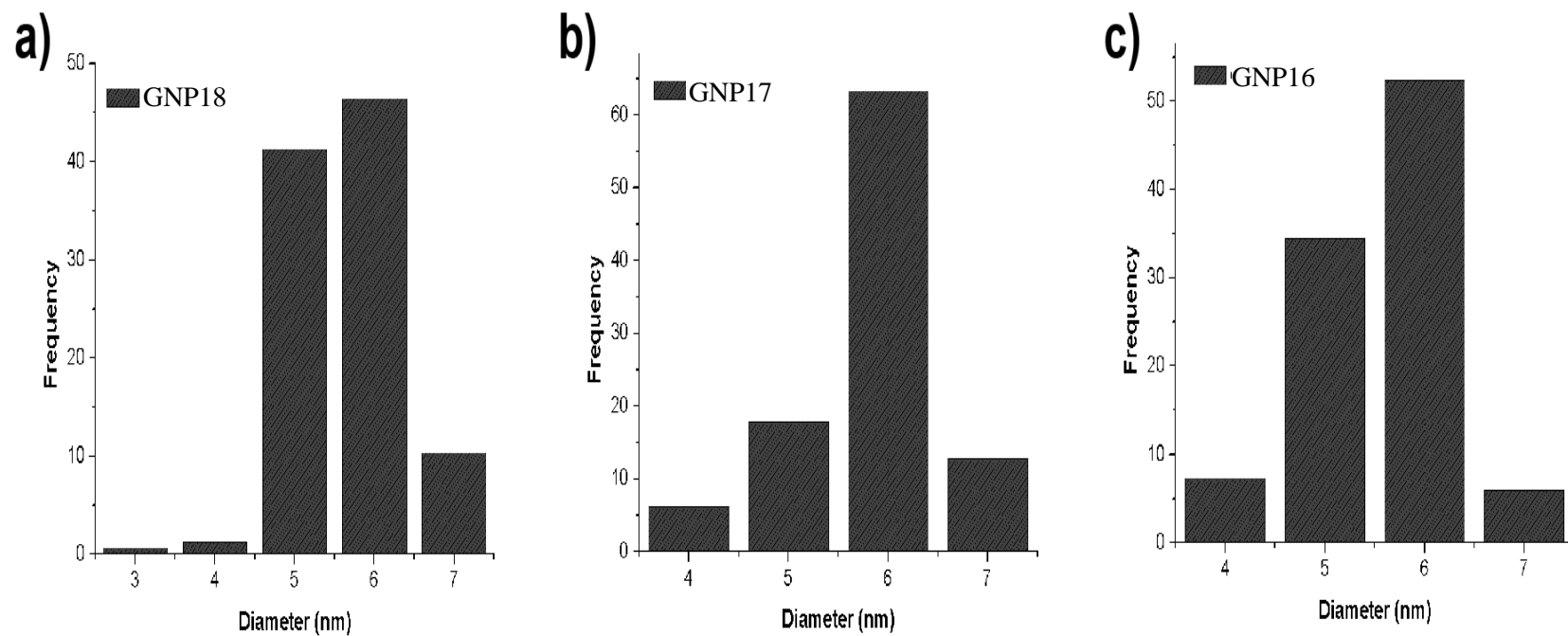


Figure 4.8. Particle size distribution (PSD) charts of TEM micrograph data for GNPs **a)** GNP18 **b)** GNP17 **c)** GNP16.

4.4.5 – SMN2 Splicing

4.4.5.1 – 18nm GNP Splicing

GNP1-9 were tested in *SMN2* splicing experiments, to identify which 2'OMe RNA surface coating would give rise to the highest exon 7 inclusion levels. GNP concentrations of 200nM, 100nM, 50nM, 10nM and 1nM were used. RNA 6 and GGA-O (Section 3.5.1.2) were used as positive controls to determine whether a multivalent system would provide greater exon 7 inclusion levels than a monovalent system. Experiments were carried out using 40% DTT-free NE, with 1.5mM rATP, 2mM MgCl₂, 20mM CrPi and 50mM KGlu. The samples were treated with proteinase K, ethanol precipitated and analysed by 6% PAGE (Figure 4.9 & Appendix Figure 5.29).

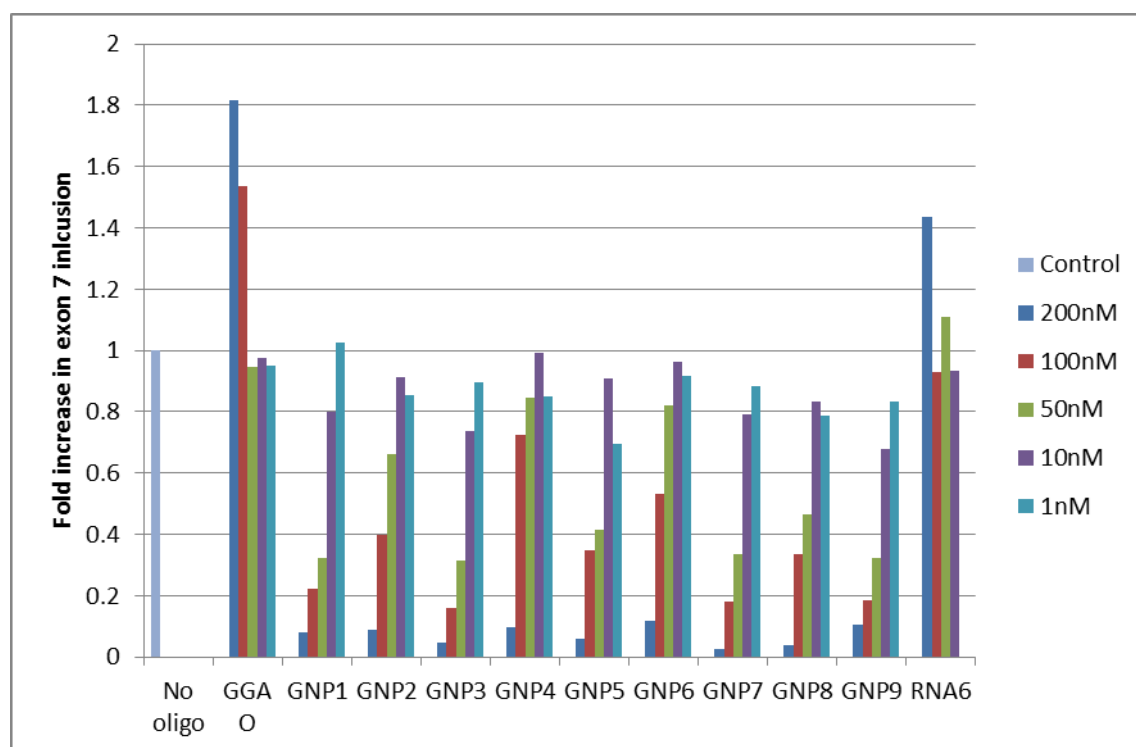


Figure 4.9. Fold increases in exon 7 inclusion in splicing reactions with 18nm GNP1-9.

GNP1-9 show inhibition of splicing at 200nM with splicing efficiency increasing with reduced nanoparticle concentration. GNP 4, which consists of an enhancer RNA17 coating, showed the greatest exon 7 inclusion with the percentage of spliced products reaching 47%, at 50nM, equalling GGA O (a 2'OMe RNA control) (Table 4.4). GNP 6, which consists of an enhancer which is conjugated through the 3' end (RNA17) and an annealer which is conjugated through the 5' end (RNA20), also showed reasonable activity with the percentage of spliced products reaching 42% for exon 7 inclusion at 50nM (Table 4.4). The only difference between GNP4 and GNP2 is the orientation in which the enhancer is bound to the GNPs. The levels of exon 7 inclusion decreased from the 14%, 42% and 47% achieved with GNP4 (when the enhancer was joined to the GNP via its 3' end) to the 4%, 24% and 34% achieved with GNP2 at 200nM, 100nM and 50nM, respectively (when the enhancer was joined to the GNP via its 5' end). This shows that the direction of conjugation is essential for enhancer activity. Comparing GNP5-8 with mixed monolayers also shows that the direction of conjugation is also essential for enhancer activity with the best mixed monolayer construct, GNP6, having the enhancers conjugated through the 3' end and the annealer conjugated through the 5' end to the GNP (Table 4.3 & Figure 4.4).

Table 4.4. Exon 7 inclusion levels of spliced products for GNP1-9, GGA-O and RNA6 (%)

Oligonucleotide Name	Exon 7 inclusion for spliced products (%)				
	200nM oligonucleotide	100nM oligonucleotide	50nM oligonucleotide	10nM oligonucleotide	1nM oligonucleotide
Water control	25	-	-	-	-
GGA O	61	49	47	46	47
GNP1	7	12	22	43	49
GNP2	4	24	34	48	56
GNP3	3	10	15	37	46
GNP4	15	42	47	53	48
GNP5	5	16	27	47	51
GNP6	11	31	42	54	46
GNP7	6	11	23	47	54
GNP8	8	18	28	49	52
GNP9	7	9	21	42	51
RNA6	51	35	47	49	49

4.4.5.2 - SMN2 Splicing with 20nm, 10nm and 5nm GNP

In *Section 4.4.4.1* it was observed that a GNP4 coating of RNA17 and GNP6 coating of RNA17 and RNA20 caused enhancement of exon 7 inclusion. Therefore, a monolayer of RNA17 and a multilayer of RNA17 and RNA20 were polyconjugated to 20nm, 10nm and 5nm GNPs (GNP10-18) using GNP concentrations of 200nM, 100nM, 50nM, 10nM, 1nM and 0.4nM. Splicing reactions were undertaken and analysed as described in *Section 4.4.4.1* (Figure 4.10 & Appendix Figure 5.30). Reactions were run against water controls of normal splicing (control) and splicing using DTT free NE dialysed control (dia control).

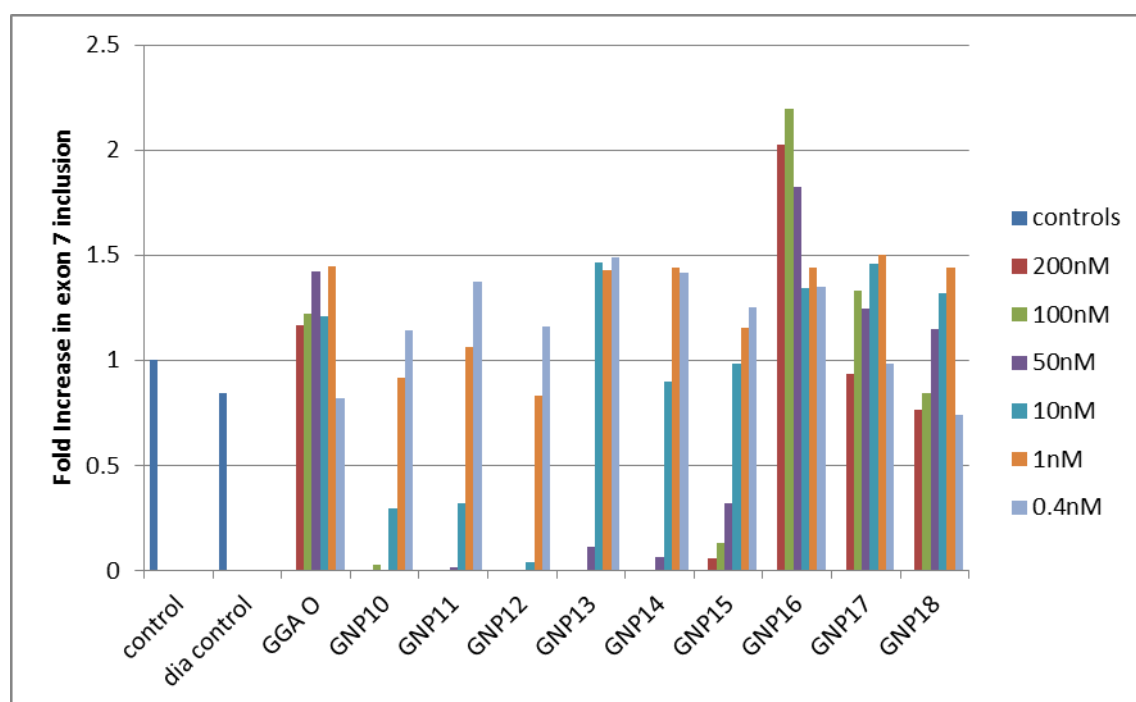


Figure 4.10. Fold increases in exon 7 inclusion in SMN2 splicing reactions using 20nm, 10nm and 5nm GNPs (GNP10-18) and GGA-O at concentrations of 200nM, 100nM, 50nM, 10nM, 1nM and 0.4nM. Where dia control is a splicing reaction using the DTT free NE and control is water control.

The data in Table 4.5 reveals that 20nm GNPs10-12 inhibit splicing at concentrations of 200nM, 100nM, 50nM and 10nM with splicing efficiency increasing as GNP concentration decreases and splicing reaching approximately basal levels at 1nM. 10nm GNPs13-15 have a similar effect with splicing inhibition occurring at concentrations of 200nM, 100nM and 50nM and stopping at 10nM (Figure 4.10). GNP13 increased exon 7 inclusion at 10nM, but the inclusion levels did not decrease to basal levels even when GNP13 concentrations were decreased, therefore it seems that even at very low concentrations the GNP13 has a positive effect on splicing (Figure 4.10). 5nm GNPs16-18 were the only nanoparticles that inhibited splicing at a similar level to GGA-O at 200nM (Table 4.5), with pre-mRNA levels going from 59% to 50% to 38% at 200nM, 100nM and 50nM, respectively compared to the basal levels of pre-mRNA of 54%. At 100nM, GNP16 significantly enhanced exon 7 inclusion levels, increasing basal splicing levels 2.2 fold (Figure 4.10). Exon 7 inclusion of spliced products increased from 43% with GGA-O to 71% with GNP16 at 200nM, revealing that in this case, the multivalent system (GNP16) gives rise to greater inclusion levels than the monovalent system (GGA-O). GNP17 only achieved percentage exon 7 inclusion levels of spliced products of 43%, 47% and 42% and GNP18 achieved 20%, 21% and 31% at 200nM, 100nM and 50nM, respectively. Even though GNP17 is not as effective at promoting exon 7 inclusion as GNP16, the levels of exon 7 inclusion achieved with GNP17 are similar to those produced by an exemplar TOES GGA O at the same concentrations ². These results therefore infer that the enhancement of exon 7 inclusion using GNP18 is due to the coating on the GNPs and not the GNPs themselves.

Table 4.5. Exon 7 inclusion and exclusion levels and the levels of pre-mRNA remaining after 2 hours of splicing reactions with GNP10-18 and GGA-O at concentrations of 200nM, 100nM, 50nM and 10nM (%). Where dia control is a splicing reaction using the DTT free NE.

GNP/ON	200nM concentration			100nM concentration			50nM concentration			10nM concentration		
	Pre-mRNA (%)	Exon 7 inclusion (%)	Exon 7 exclusion (%)	Pre-mRNA (%)	Exon 7 inclusion (%)	Exon 7 exclusion (%)	Pre-mRNA (%)	Exon 7 inclusion (%)	Exon 7 exclusion (%)	Pre-mRNA (%)	Exon 7 inclusion (%)	Exon 7 exclusion (%)
Water Control	37	15	47	30	17	53	38	11	51	-	-	-
Dia control	54	17	29	54	11	35	56	8	36	-	-	-
GGA-O	61	17	23	60	18	22	52	21	27	50	17	33
GNP10	100	0	0	96	0	4	100	0	0	61	4	35
GNP11	100	0	0	100	0	0	100	0	0	50	5	46
GNP12	100	0	0	100	0	0	100	0	0	76	1	23
GNP13	100	0	0	100	0	0	97	2	2	39	21	40
GNP14	100	0	0	100	0	0	32	1	7	57	13	31
GNP15	93	1	6	54	2	44	48	5	48	47	14	39
GNP16	59	29	12	50	32	18	38	26	36	42	19	38
GNP17	68	14	18	59	19	22	57	18	25	51	21	28
GNP18	45	11	44	41	12	46	46	17	37	46	35	35

4.4.5.3 – Triplicate GNP16-18 *SMN2* Splicing

5nm GNP15-18 constructs were then run in triplicate by Rachel Dickinson to ensure the data above was reproducible. This assay series was conducted and analysed as described in *Section 4.4.4.1* (Figure 4.11 & Appendix Figure 5.31).

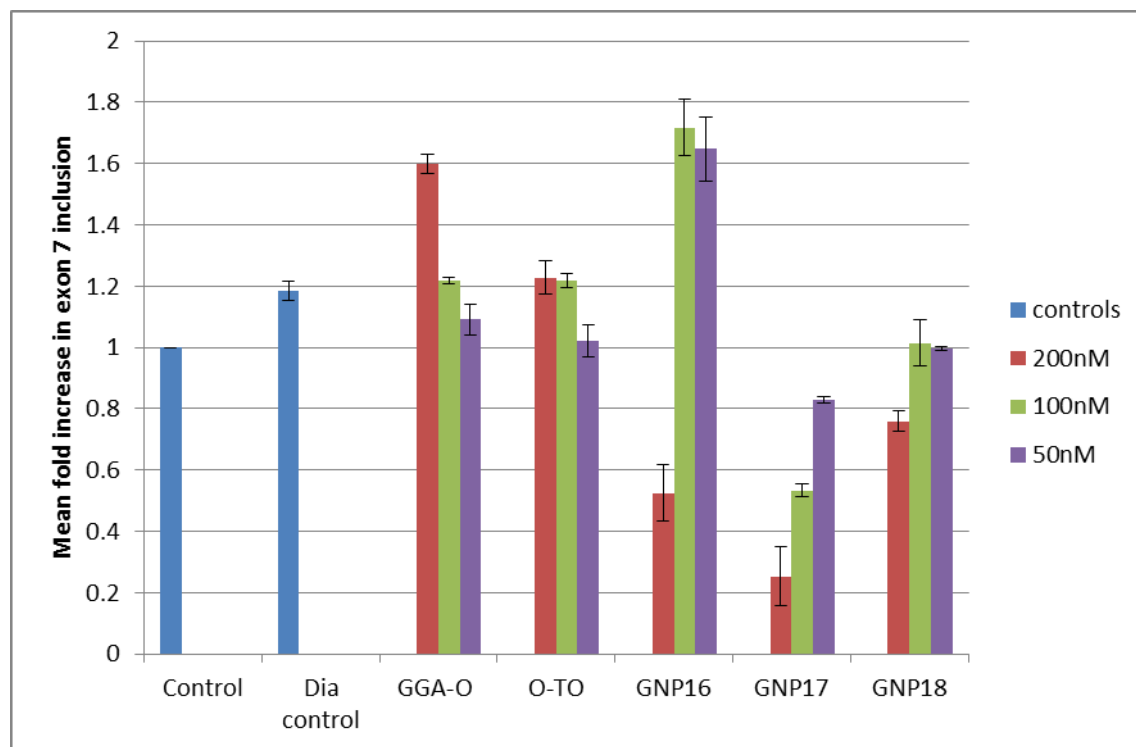


Figure 4.11. Mean increases in exon 7 inclusion (depicted as a fold increase in Exon 7 inclusion) of *SMN2* pre-mRNA for 5nm GNP15-18, GGA-O and O-TO. Where GGA O is a 2'OMe TOES and O-TO is the 2'OMe enhancer sequence. Dia control is a splicing reaction using the DTT free NE.

Figure 4.11 reveals that GNP16 with a RNA17 monolayer coating, GNP17 with an RNA17 and RNA20 mixed monolayer and GNP18 which has a PEG-thiol coating, all had an inhibitory effect on splicing at 200nM. Such an effect was not seen in Figure 4.10 with these particles. This is thought to be due the fact that a different batch of NE was used in the reactions in Figure 4.10, which could have affected the splicing results¹³¹. There was a 1.6 fold increase in

exon 7 inclusion between the monovalent system O-TO, which is just an enhancer sequence, and GNP16 which is a multivalent system, revealing the benefits of using a multivalent system.

Table 4.6. Mean exon 7 inclusion of spliced products for GNP16-18, GGA-O and O-TO at 200nM, 100nM and 50nM. Where dia control is a splicing reaction using the DTT free NE.

GNP/ON	Mean exon 7 inclusion of spliced products (%)		
	200nM GNP/ON	100nM GNP/ON	50nM GNP/ON
Control	16	-	-
Dia control	44	-	-
GGA-O	67	53	42
O-TO	50	47	38
GNP16	80	75	63
GNP17	29	40	47
GNP18	39	43	42

Although GNP16 caused a decrease in splicing levels at 200nM (Figure 4.11) it gave rise to exon 7 inclusion levels of 80%, whereas GGA-O only achieved 67% exon 7 inclusion at 200nM (Table 4.5). GNP16 gave rise to higher levels of exon 7 inclusion than GNP17, GNP18 and GGA-O at all concentrations tested. Therefore the RNA17 monolayer on GNP16 is more effective at increasing exon 7 inclusion levels in *SMN2* than the RNA17 and RNA20 mixed monolayer of GNP17 and the PEG-thiol monolayer of GNP18. This reveals that best GNP coating to use for enhancing exon 7 inclusion is an RNA17 monolayer.

4.4.6 – GNP16-18 in Adenovirus Splicing

Since GNP16 contains no annealing sequence, it was thought that GNP16 would be an effective splicing enhancer for other genes. To test this hypothesis two adenovirus constructs were chosen, A2 and A3. A2 is a mini gene with two exons with an u/s and a d/s splice site in exon 1. A3 is the same mini gene but contains an enhancer sequence at the 5' end of the transcript (Figure 4.12).

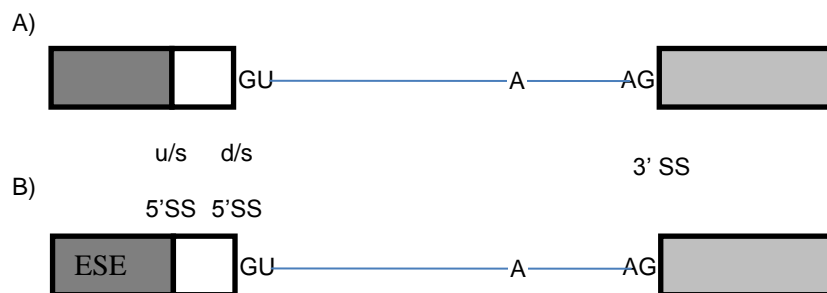
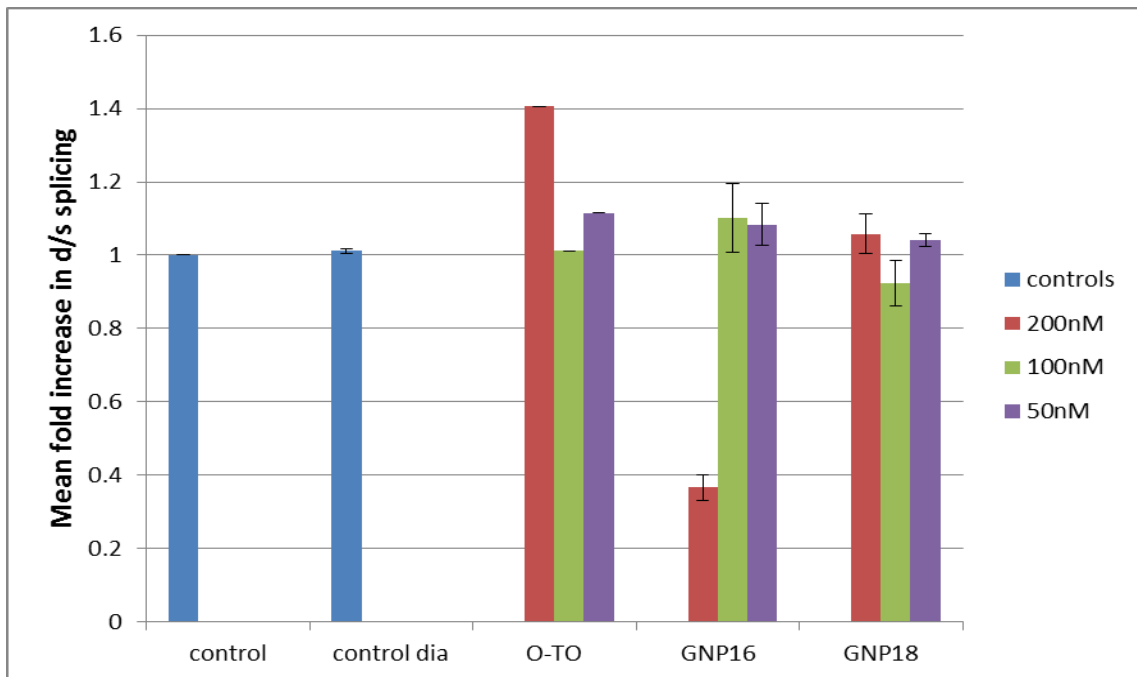


Figure 4.12. Adenovirus mini genes used for testing splicing enhancement by GNP16 and GNP18. A) A2 adenovirus construct. B) A3 adenovirus construct with an ESE at the 5' end of the transcript.

Splicing reactions were conducted by Rachel Dickinson, with DTT-free NE for 1 hour at 30°C. The reactions were then treated with proteinase K and ethanol precipitated. The products were analysed by 6% denaturing PAGE (Figure 4.13 & Appendix Figure 5.32).

A)



B)

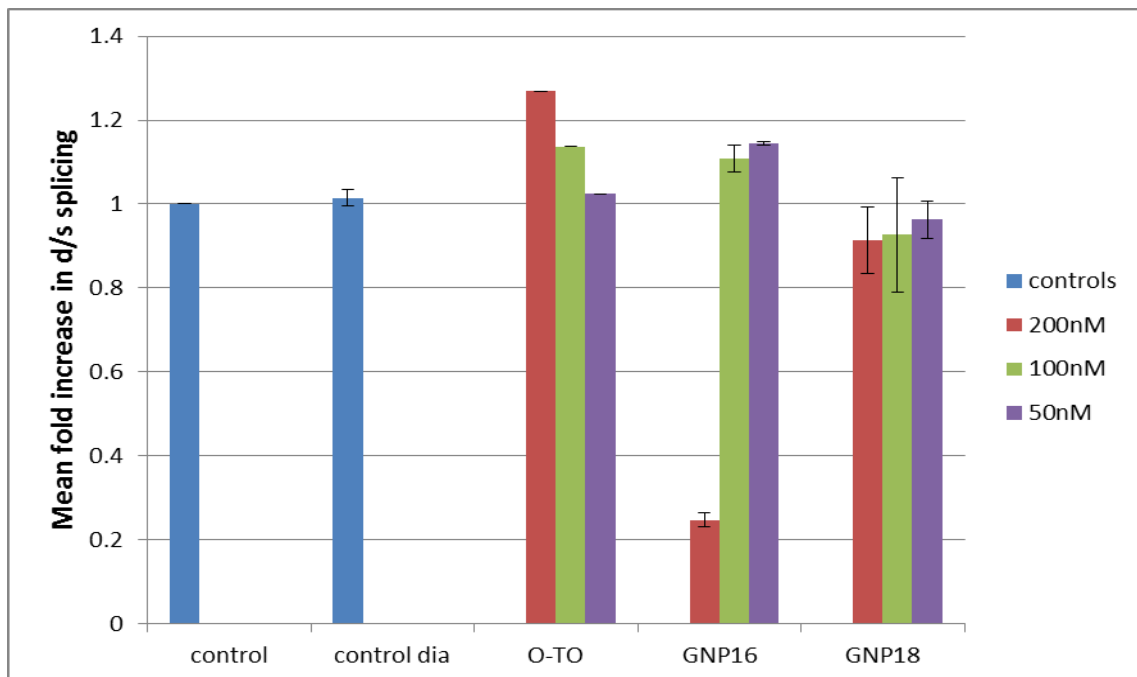


Figure 4.13. Mean fold increases in Adenovirus D/S splicing for GNP16, GNP18 and O-TO in A) construct A2 and B) construct A3. Where dia control is a splicing reaction using the DTT free NE. Splicing was carried out by Rachel Dickinson.

Table 4.7. Mean percentage of Adenovirus A2 and A3 spliced to form d/s spliced products with GNP16, GNP18 and O-TO. Where dia control is a splicing reaction using the DTT free NE

GNP/ON	Mean percentage of A2 spliced to form d/s product			Mean percentage of A3 spliced to form d/s product		
	200nM GNP/ ON	100nM GNP/ ON	50nM GNP/ ON	200nM GNP/ ON	100nM GNP/ ON	50nM GNP/ ON
Control	44	-	-	21	-	-
Dia control	58	-	-	25	-	-
O-TO	64	45	56	32	26	23
GNP16	72	76	65	50	40	30
GNP18	62	51	60	27	25	25

Figure 4.12 reveals that splicing of both A2 and A3 was inhibited when a concentration of 200nM of GNP16 was used; however, GNP18 did not inhibit splicing. The inhibition of splicing is therefore due to the RNA17 coating on the GNP16, not the GNP itself as otherwise both GNP16 and GNP18 would inhibit splicing. The multivalent GNP16 system achieved 72% d/s splicing of A2 at 200nM, whereas the monovalent enhancer version, O-TO, only achieved 64% (Table 4.7). A3 splicing data was similar to A2 splicing data, with GNP16 achieving 50% d/s splicing of A3 at 200nM compared to 32% for O-TO (Table 4.7). GNP18 in both A2 and A3 only achieved basal splicing levels revealing that the splicing enhancing effect of GNP16 was due to the GNP coating and not the GNPs themselves.

4.4.7 – Discussion

The effectiveness of a multivalent system depends on two main things: particle size and surface coating. With particle sizes of 20nm, 18nm and 10nm, at high concentrations, GNP surface coatings had less of an effect on exon 7 inclusion levels than at lower concentrations. This could be due to the GNPs blocking key protein-RNA and RNA-RNA interactions required for splicing, due to their comparatively large size (*Section 4.1*). However, at lower concentrations, the RNA17 monolayer coating on GNP4 and mixed monolayer coating of RNA17 and RNA20 on GNP6 did give rise to an increase in exon 7 inclusion levels at GNP concentrations of 100nM and 50nM (*Section 4.4.5.1*, Table 4.4). The directionality of conjugation also plays a defining role in enhancer activity as shown by GNP2 and GNP4 which have the same RNA sequence, which is conjugated through the 5' end in GNP2 and through the 3' end in GNP4 (*Section 4.4.3*, Table 4.3). This difference in the directionality of the conjugation has a dramatic effect on exon 7 inclusion with GNP4 giving rise to 14%, 42% and 47% and GNP2 only giving rise to 4%, 24% and 34% exon 7 inclusion of spliced products at GNP concentrations of 200nM, 100nM and 50nM, respectively (*Section 4.4.5.1*, Table 4.4). GNP5-8, which all have mixed monolayer surface coatings, also show that the directionality of conjugation is important for enhancing exon 7 inclusion, with the best mixed monolayer coating of RNA17 and RNA 20 on GNP6 having the enhancers conjugated through the 3' end and the annealer conjugated through the 5' end to the GNP (*Section 4.4.5.1*, Table 4.4). This indicates that particle size, surface coating and directionality of conjugation are all important for splicing enhancement.

As specified in *Section 4.4.5.4*, 20nm and 10nm GNPs inhibit splicing with these inhibitory effects being concentration dependent and negligible by 1nM and 10nM respectively (*Section 4.4.5.4*, Table 4.5). 5nm particles (GNP16-18) did show inhibition of splicing at 200nM, however the percentage of exon 7 inclusion still remained high (*Section 4.4.5.5*, Table 4.6).

GNP16, with only an enhancer coating conjugated to the GNP through the 3' end, had the greatest effect on splicing, with the exon 7 inclusion percentage of spliced products reaching 80% compared to only 50% with O-TO (enhancer only) at 200nM (*Section 4.4.5.5*, Table 4.6). This shows that a multivalent system enhances exon 7 inclusion far more than a monovalent counterpart. GNP17, with both an annealer and an enhancer sequence, and GNP18, with a PEG coating, did not enhance splicing as effectively as GNP16, with exon 7 inclusion of spliced products reaching 29% and 39% for GNP17 and GNP18 respectively at 200nM (*Section 4.4.5.5*, Table 4.6). This infers that with 5nm GNPs, the coating is essential for enhancer activity and that a higher effective concentration of the enhancer on the surface leads to increased exon 7 inclusion. The low levels of exon 7 inclusion that occurred with the mixed monolayer coated GNP17 was surprising as it was thought that being able to bind the pre-mRNA via its annealer sequences, GNP17 would give rise to greater exon 7 inclusion levels. The same effect was seen with the 18nm GNPs (*Section 4.4.5.1*, Table 4.4) where GNP4, which has no annealer sequences, provided greater levels of exon 7 inclusion than GNP6 which does have annealer sequences. It was originally proposed that the levels of exon 7 inclusion may have been low due to GNP6 and GNP17 binding the pre-mRNA strongly which has been shown to decrease exon 7 inclusion levels (*Section*

2.1.1)². This however was ruled out when a pull down experiment revealed that there was only very weak, if any, binding occurring between GNP6 and pre-mRNA (Section 4.4.5.2, Figure 4.9). It was then thought that GNP16 and GNP4 may have given rise to such higher exon 7 inclusion levels than GNP17 and GNP6 respectively, due to having higher numbers of enhancer sequences on their surfaces.

GNP16 does not have any annealing sequences to direct it to a specific region of pre-mRNA, therefore it stands to reason that GNP16 could be used to manipulate splicing in other genes, as well as *SMN2*. GNP16 increased D/S splicing in both Adenovirus mini genes *A2* and *A3*, with the greatest increase being achieved with 200nM GNP16 (Section 4.4.6, Table 4.7). Therefore it is concluded that GNP conjugates containing enhancer sequences, in general, act as enhancers for alternative RNA splicing.

4.4.8 – Conclusion

Multivalent system GNP16 is more effective at increasing exon 7 inclusion levels in *SNM2* and Andenovirus D/S splicing levels of *A2* and *A3* than monovalent system TOES. Therefore it is highly probable that GNP16 could be used as a general method for inducing splicing. GNP16, with an RNA17 monolayer coating, was the most effective splicing enhancer of all those produced. GNP16 works effectively as a general enhancer, but attaching annealer sequences onto it, significantly reduced its overall activity. This problem would need to be overcome in order to enable GNP16 to bind specifically to pre-mRNA, without reducing its activity.

4.5 – Experimental

20nm, 10nm and 5nm GNP were purchased from BBI international. 18nm particles were prepared by Dr Vanessa Bonnard according to the method described by S Kumar *et al*¹⁸⁹. Compound **9** (Section 4.5.2.2) was prepared by Dr Vanessa Bonnard¹⁹³.

4.5.1 – General Protocols

4.5.1.1 – Standard Ligand Exchange Protocol

BSPP (6mg, 1.12×10^{-5} mol) was added to an aqueous solution of citrate coated GNPs (10ml) covered with foil and left at stirring at 200rpm overnight at 25°C. Sodium chloride (2mg, 3.41×10^{-5} mol) was then progressively added to the BSPP coated GNPs. After each 2mg addition of NaCl the sample was vortexed until the solution turned purple. The GNP solution was split into 1ml aliquots and centrifuged at 12,000rpm for 10 minutes at room temperature. The supernatant was removed from each aliquot and the GNP combined in a BSPP/water solution (200µl, 2.5mM). Methanol (200µl) was then added, vortexed and centrifuged again at 12,000rpm for 10 minutes at room temperature. The supernatant was removed again and the particles were re-suspended in an aqueous solution of BSPP solution (200µl, 2.5mM) and stored at 5°C.

4.5.1.2 – Preparation of 2'OMe RNA GNP Conjugates

TCEP was dissolved in 0.5 x TBE (pH = 8.3) to make a 20mM solution.

To a solution of 2'OMe RNA17-20 (6µl, 700µM) was added TCEP (69µl, 20mM, pH 8.3). The reaction was vortexed and left to shake at 750 rpm for 1 hour at 25°C.

To a solution of GNP-BSP (50 μ l) was then added an aqueous solution of 2'OMe RNA lipoic acid (22.9 μ l, 56 μ M, 200eq) and 0.5 x TBE, 0.015% SDS (300 μ l). The samples were vortexed and then shaken at 650 rpm at 25°C for 20min. A solution of 2M NaCl (Table 4.8) was then added every 20 minutes to gradually increase the salt concentration from 50mM to 700mM. After each salt addition the samples were vortexed, sonicated for 10 seconds, vortexed again and then returned to the shaker for 20min, 750rpm at 25°C. If the samples turned purple (a sign of aggregation¹⁸⁷) before the 700mM NaCl addition then 0.5x TBE, 0.015% SDS was added until the solution turned red again and left overnight at 25°C, 750rpm. For samples that did not aggregate after the 700mM addition of NaCl, these samples were left for a further 20 minutes to check to see whether they would aggregate (turn purple¹⁸⁷) and if they were still red they were left overnight at 25°C, 750rpm.

After the overnight incubation, the reactions were vortexed and then centrifuged for 10 minutes at 10°C, 13,200rpm. The supernatants were removed and the GNPs were re-suspended in 0.02% SDS in ddH₂O (300 μ l). The samples were vortexed again followed by centrifugation at 10°C for 10 minutes at 13,200rpm. This process was repeated four times; re-suspending each time in 0.02% SDS in ddH₂O (300 μ l). After the GNPs had been washed four times the samples were centrifuged again at 13,200rpm, 10°C for 10 minutes. The supernatants were discarded and the samples re-dissolved in 0.02% SDS in ddH₂O (50 μ l).

Table 4.8. Shows the volume of 2M NaCl added to achieve salt concentrations for the preparation of GNP1-18. NaCl concentrations ranged from 50mM to 700mM

Salt concentration required	Volume of 2M NaCl in 0.5x TBE, 0.015% SDS (µl)	New total volume (µl)
50mM	9.6	382
100mM	9.8	392
200mM	20.6	412
300mM	21.7	434
400mM	22.9	457
500mM	24.1	481
600mM	25.4	506
700mM	26.7	533

4.5.1.3 – ³²P Labelled (Hot) Transcription Protocol

Hot transcription buffer - Tris (0.2ml, 1M, pH = 7.5), MgCl₂ (30µl, 1M), spermidine. HCl (11.5µl, 0.87M), NaCl (10µl, 5M) and water (0.75ml).

Low GTP rNTP mix - rATP (5µl, 100mM), rCTP (5µl, 100mM), rUTP (5µl, 100mM) and rGTP (0.5µl, 100mM).

A standard transcription reaction comprised the following:

2µl Hot transcription buffer

2.5µl water

1µl Cap

1µl low GTP rNTP mix

1µl DNA template (50ng/µl from PCR)

0.5µl 0.1M DTT

These reagents were combined, vortexed to ensure thorough mixing before RNaseOUT™ (0.5µl) and T7 polymerase (0.5µl) was added. The reaction was vortexed again before finally adding ³²P labelled GTP (1µl). The reaction was mixed with a pipette and then placed in a water bath at 37°C for 1 hour.

After incubation for one hour the sample was treated with 10µl formamide dyes and then run on a 6% denaturing polyacrylamide gel at 18W for 2½ hours. The glass plates were separated and saran wrap was placed over the open side of the gel. Glow in the dark stickers placed on to the saran wrap to observe the reference points on the developed x-ray film. The gel was then exposed to an x-ray film in a dark room for 1 – 3 minutes before developing the film. The developed film was then lined up using the reference stickers and the band excised from the gel. The excised band was treated with RNA elution buffer (350µl) and left at 5°C overnight. After incubation the sample was heated to 30°C for 10 minutes to re-dissolve the precipitated SDS before transferring the supernatant to a new tube. Ethanol (1.05ml, 3x vol of RNA elution buffer) was then added to the supernatant, vortexed and then centrifuged for 30 minutes at 13,400rpm. The supernatant was removed and the RNA pellet washed with ethanol (250µl). The mixture was centrifuged for 15 minutes at 13,400rpm. The supernatant was once again removed and the pellet dried and re-suspended in water.

4.5.1.4 – GNP *in vitro* SMN2 Splicing

NE was dialysed for 2 hours in the cold room with DTT free buffer D. DTT free buffer D was made up from:-

800µl 1M HEPES buffer pH 8 (20mM final concentration)

5ml 80% Glycerol (10% final concentration)

16µl 0.5M EDTA (0.2mM final concentration)

4ml 1M KCl (100mM final concentration)

30.19ml water

During the dialysis, relevant stocks were prepared from the 1µM GNP solution to enable the desired concentration to be tested in splicing reaction. 1µl aliquots of these stocks were then treated with ³²P-labelled RNA (0.5µl) and heated in a 30°C oven for 10 minutes. After incubation the samples were placed on ice for 10 minutes. While the samples were on ice a master mix was prepared. The master mix consisted of:

6.1µl 100mM ATP

15µl 0.5M CrPi

9.5µl 80mM MgCl₂

3.8µl 1M HEPES KOH pH 7.5

3.8µl 1% NP40

75µl Kglu

Once the dialysis had finished the dialysed NE (150µl) was added to the master mix, vortexed and then 3.5µl added to each reaction. The reactions were vortexed before being placed in a 30°C oven for the duration of the splicing reaction. Once the splicing reaction was complete proteinase K (10µl, 1 in 25 dilutions in PK buffer) was added to each of the reactions and heated to 37°C

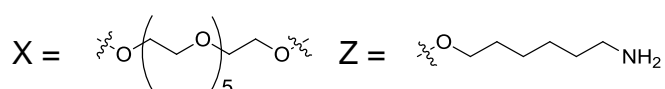
for 30 minutes. After 30 minutes the samples were treated with ethanol (200µl) and centrifuged at 6,200rpm (microtitre plate centrifuge) for 15min. The supernatants were then removed and the pellets washed with ethanol (200µl) and centrifuged again for 5 minutes. The supernatants were again removed and the pellets dried in a vacuum desiccator. The samples were then treated with 10ul formamide dyes before being either placed in the -80°C freezer or used directly. The products were analysed by running a 6% denaturing polyacrylamide gel at 30W for 1³/₄ hours. The plates were then separated and the gel transferred to 3MTM paper. The open side of the gel was covered in saran wrap before the gel was dried on a gel drier for approximately an hour. Once dried the gel was exposed to a phosphor screen overnight in a cassette. In the morning the phosphor screen was imaged using a Packard CycloneTM and the data quantified.

4.5.2 – Preparation of 2'OMe RNA ONs for GNPs.

4.5.2.1 – 2'OMe RNA Solid Phase Synthesis

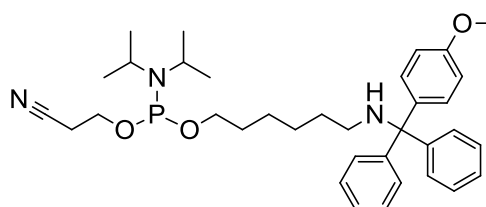
Table 4.9. RNA ONs prepared by solid phase synthesis. Building blocks X and Z are shown below

Sequence Name	Sequence – 5' to 3'
RNA1_I	AGGAGGACGGAGGACGGAGGACAXXXZ
RNA2_D	GAUUUUGUCUAAAACXXXZ
RNA1_J	ZXXXAGGAGGACGGAGGACGGAGGACA
RNA2_E	ZXXXGAUUUUGUCUAAAAC



2'OMe RNA1_I, RNA1_J, RNA2_D and RNA2_E were prepared by solid phase synthesis using the DMT- and β -(cyanoethyl) phosphoramidite method on 1 μ mol CPG supports by an Applied Biosystems 394 synthesizer. A double coupling protocol and extension of coupling times to 18 minutes was applied for ONs RNA1_I and RNA1_J, with coupling times for the other strands at 11 minutes (Table 4.9).

ONs RNA1_I and RNA2_D were synthesised with an MMT¹⁹⁴ (Figure 4.14) protected amine which was removed manually as the MMT protecting group required extended deprotection times.



MMT protected amino modified phosphoramidite

Figure 4.14. Structure of 5' amino modifier with MMT protecting group used in solid phase synthesis.

ON were cleaved from the solid support and deprotected by aminolysis (35% aqueous ammonia) at 55°C overnight. The aqueous ammonia was removed by evaporation and the samples worked up using an Illustra NAPTM 25 columns, prior to lyophilisation and re-suspension in water (

Table 4.10).

Table 4.10. Summary of RNA1_I, RNA1_J, RNA2_D and RNA2_E prepared by solid phase synthesis. Mass spectrometric analysis included in appendix (Figure 5.27)

Sequence	Progress	Amount (mg)	Expect MW	Mass Spec MW
RNA1_I	Desalted	3.79	9163.62	9163.49
RNA2_D	Desalted	3.61	6143.60	6144.14
RNA1_J	Desalted	4.95	9163.62	9163.60
RNA2_E	Desalted	3.44	6143.60	6143.10

4.5.2.2 – Preparation of RNA Lipoic Acid Derivatives

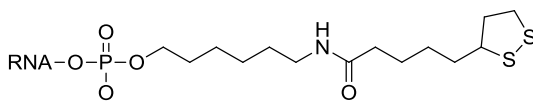


Table 4.11. Products formed from RNA1_I, RNA1_J, RNA2_D and RNA2_E reaction with compound **18**.

Product Name	Starting material
RNA17	RNA1_I
RNA18	RNA2_D
RNA19	RNA1_J
RNA20	RNA2_E

RNA1_I (500µl, 414µM), RNA2_D (500µl, 587µM), RNA1_J (500µl, 540µM) and RNA2_E (500µl, 560µM) were evaporated to dryness, in separate tubes, before being re-dissolved in sodium hydrogen carbonate (180µl, 0.025M, pH 8). Lipoic

acid NHS (**18**) (40mg, 1.32×10^{-4} mol) was dissolved in DNA grade acetonitrile (1.68ml, 78.6mM). A proportion of compound **18** would not dissolve. This solid was assumed to be hydrolysed lipoic acid. In this case, the sample was centrifuged and the supernatant transferred to a new tube. 420µl of the lipoic acid-NHS solution was then added to each of the 2'OMe RNA, vortexed and then left overnight at 25°C mixing at 300rpm. The product was concentrated to remove the acetonitrile, desalted through a GE healthcare NAP 25 column before being lyophilised and re-dissolved in pure water (Table 4.12).

Table 4.12. Summary of RNA17-20 prepared by NHS coupling of compound **18** to RNA1_I, RNA1_J, RNA2_D and RNA2_E. Mass spectrometric analysis included in appendix (Figure 5.28)

Sequence	Progress	Amount (mg)	Expect MW	MALDI-TOF MW
RNA17	Desalted	1.33	9351.63	9350.90
RNA18	Desalted	1.43	6331.61	6331.12
RNA19	Desalted	1.73	9351.63	9352.73
RNA20	Desalted	1.42	6331.61	6331.18

4.5.3 – GNP Preparation and Polyconjugation

4.5.3.1 – Ligand Exchange of 18nm GNP-Ci

Citrate coated 18nm GNP (10ml, batch no: 11.03.11) were treated with BSPP (6mg, 1.12×10^{-5} mol) and the reaction carried out as described in the standard ligand exchange protocol (Section 4.5.1.1). The concentration of the nanoparticle solution was determined from the 18nm absorption coefficient

$$\epsilon = 6.64 \times 10^8 \text{ at } 524\text{nm}^{195}$$

Concentration = 116nM in 200µl

4.5.3.2 – Polyconjugation of 18nm GNP-BSPP with 2'OMe RNA

TCEP (3mg, 1.04×10^{-5}) was dissolved in 0.5x TBE (523.3µl, pH = 8.3) to make a 20mM solution. TCEP (69µl, 20mM) was then added to aqueous 2'OMe RNA17-20 and peg-SH (6µl, 700µM), vortexed and left at 25°C for 1 hour mixing at 750rpm.

18nm GNP-BSPP (50µl, 128nM) was aliquoted into nine tubes. A solution of reduced aqueous solution of a 2'OMe RNA ON (Table 4.13) in 0.5 x TBE, 0.015% SDS (300µl) was added to the GNP-BSPP solution, vortexed and then left at 25°C, 750rpm for 20 minutes. The reaction was then carried out as in the standard polyconjugation protocol (*Section 4.5.1.2*) (Table 4.15 indicating the volumes of salt solution added). GNP conjugates GNP3, GNP8 and GNP9 had turned a pale red colour followed by the formation of a black precipitate (aggregates¹⁸⁷) before the salt concentration had reached 700mM. In this case these reactions were treated with 0.5x TBE, 0.015% SDS, until they turned red again, before being left overnight at 25°C, 750rpm. In the morning the particles were work up as in the standard polyconjugation protocol. GNP4 and GNP6 show aggregated particles (black precipitate¹⁸⁷) even after the four washes and so the reactions were centrifuged for 3 seconds to precipitate the aggregated particles. The supernatant was transferred to a new tube and the standard polyconjugation work up continued. The aggregated particles were treated with 0.02% SDS in ddH₂O (50µl), sonicated to break up the aggregates, and kept separate. Polyconjugated particles were examined by 2% agarose gel at 100V for 1 hour (*Section, 4.4.3, Figure 4.5*).

Table 4.13. RNA17-20 polyconjugated onto GNP1-9 used in this study

Reference	Particle Size	Polyconjugate coating	Volume of reduced aqueous 2'OMe RNA ON (56µM) added to NP (µl)
GNP1	18nm	RNA20	22.9
GNP2	18nm	RNA19	22.9
GNP3	18nm	RNA18	22.9
GNP4	18nm	RNA17	22.9
GNP5	18nm	RNA19	11.5
		RNA20	11.5
GNP6	18nm	RNA17	11.5
		RNA20	11.5
GNP7	18nm	RNA18	11.5
		RNA19	11.5
GNP8	18nm	RNA17	11.5
		RNA18	11.5
GNP9	18nm	peg-SH	22.9

The amount of polyconjugated material recovered after precipitations is reported in Table 4.15.

Table 4.14. Volumes of 2M NaCl in 0.5x TBE, 0.015% SDS added to achieve the final concentrations used in splicing assays described in (Section 4.5.1.2)

Salt concentration required	Volume of 2M NaCl in 0.5x TBE, 0.015% SDS (µl)	New total volume (µl)
50mM	9.6	382
100mM	9.8	392
200mM	20.6	413
300mM	21.7	435
400mM	22.9	458
500mM	24.1	482
600mM	25.4	507
700mM	26.7	534

Table 4.15. Concentration of NP conjugates GNP1-9 after removing excess RNA strands

Sequence	Concentration (nM)	Volume (μl)
GNP1	105	50
GNP2	113	50
GNP3	103.5	50
GNP4	109	50
GNP5	109.3	50
GNP6	85.3	50
GNP7	109.8	50
GNP8	100.5	50
GNP9	97.3	50

4.5.3.3 – Preparation of BSPP Coated 20nm, 10nm and 5nm Particles

10nm (10ml) and 5nm (10ml) citrate coated GNP were each treated with BSPP (6mg, 1.12×10^{-5} mol) in separate tubes and the experiment carried out as in the standard ligand exchange protocol (Section 4.5.1.1). 20nm (20ml) citrate coated GNP were also treated with BSPP (12mg, 2.25×10^{-5} mol) and the experiment carried out as described in the standard ligand exchange protocol (Section 4.5.1.1).

The concentration of the nanoparticle solution was determined from the absorption coefficient reported by BBI international with 20nm particles $\epsilon = 9.406 \times 10^8$, 10nm particles $\epsilon = 9.550 \times 10^7$ and 5nm particles $\epsilon = 9.696 \times 10^6$ at 524nm¹⁹⁵.

20nm GNP = 129.7nM in 200µl

10nm GNP = 467nM in 200µl

5nm GNP = 3.48µM in 200ul

4.5.3.4 – Preparation of Polyconjugated 20nm GNP

Table 4.16. Polyconjugated GNP10-12 used in splicing analyses

(Section 4.5.1.2)

Reference	Particle Size	Strands conjugated	Volume of reduced aqueous 2'OMe RNA ON (56µM) added to NP (µl)
GNP10	20nm	RNA17	23.2
GNP11	20nm	RNA17	11.6
		RNA20	11.6
GNP12	20nm	peg-SH	23.2

TCEP (3.3mg, 1.15×10^{-5} mol) was dissolved in 0.5x TBE (575.6µl, pH = 8.3) to make a 20mM solution. TCEP (115µl, 20mM, 2.3µmol) was then added to the aqueous 2'OMe RNA17, 2'OMe RNA20 and peg-SH (10µl, 700µM), vortexed and left at 25°C for 1 hour mixing at 750rpm.

20nm GNP-BSPP (50µl, 129.7nM) was aliquoted into three tubes. A solution of reduced aqueous 2'OMe RNA ON (Table 4.16 in 0.5 TBE, 0.015% SDS (300µl) was added to the GNP-BSPP, vortexed and then left at 25°C, 750rpm for 20 minutes. The reaction was then carried out as in the standard polyconjugation protocol (see Table 4.17 for the volume of salt solution added).

Table 4.17. Volumes of 2M NaCl in 0.5x TBE, 0.015% SDS added to achieve the final concentrations used in splicing assays described in (Section 4.5.1.2)

Salt concentration required	Volume of 2M NaCl in 0.5x TBE, 0.015% SDS (μl)	New total volume (μl)
50mM	9.6	383
100mM	9.8	393
200mM	20.7	413
300mM	21.8	435
400mM	22.9	458
500mM	24.1	482
600mM	25.4	507
700mM	26.7	534

Particles GNP12 aggregated before the salt concentration had reached 700mM and so was treated with 0.5x TBE, 0.015% SDS, until the solution turned red again, before being left overnight at 25°C, 750rpm. In the morning the particles were work up as in the standard polyconjugation protocol (Section 4.5.1.2).

Polyconjugated particles were examined by 2% agarose gel at 100V for 1 hour (Section 4.4.3, Figure 4.5). All the particles concentrations were determined and the concentration adjusted so that each solution had a nanoparticle concentration of 1μM for the splicing reactions.

4.5.3.5 – Preparation of Polyconjugated 10nm GNPs

TCEP (3.6mg, 1.26×10^{-5} mol) was dissolved in 0.5x TBE (418.6μl, pH = 8.3) to make a 30mM solution. TCEP (45μl, 30mM, 1.35μmol) was then added to the

aqueous 2'OMe RNA17, 2'OMe RNA20 and peg-SH (15µl, 700µM), vortexed and left at 25°C for 1 hour mixing at 750rpm.

Table 4.18. RNA ONs RNA18 and RNA19 used in the preparation of GNP13-15 and the volume of reduced 2'OMe RNA ON added to the GNP solution

Reference	Particle Size	Strands Conjugated	Volume of reduced aqueous 2'OMe RNA ON (175µM) added to NP (µl)
GNP13	10nm	RNA17	18.7
GNP14	10nm	RNA17	9.4
		RNA20	9.4
GNP15	10nm	peg-SH	18.7

Table 4.19. Volumes of 2M NaCl in 0.5x TBE, 0.015% SDS added to achieve the final concentrations used in splicing assays described in (Section 4.5.1.2)

Salt concentration required	Volume of 2M NaCl in 0.5x TBE, 0.015% SDS (µl)	New total volume (µl)
50mM	9.5	378
100mM	9.7	388
200mM	20.4	408
300mM	21.5	430
400mM	22.6	452
500mM	23.8	476
600mM	25.1	501
700mM	26.4	528

10nm GNP-BSPP (50µl, 328nM) was aliquoted out into three tubes. A solution of reduced aqueous 2'OMe RNA strands (Table 4.18) in 0.5 x TBE, 0.015% SDS (300µl) was added to the GNP-BSPP solutions, vortexed and then left at 25°C, 750rpm for 20 minutes. The reaction was then carried out as in the standard polyconjugation protocol (Table 4.19). Polyconjugated particles were examined by 2% agarose gel at 100V for 1 hour (Section 4.4.3, Figure 4.5). All the particles concentrations were determined and the concentration adjusted to 1µM for the splicing reactions.

4.5.3.6 – Preparation of Polyconjugated 5nm LA-RNA GNP

TCEP (3.6mg, 1.26×10^{-5} mol) was dissolved in 0.5 x TBE (48.3µl, pH = 8.3) to make a 260mM solution. TCEP (5µl, 260mM, 1.3µmol) was then added to the aqueous 2'OMe RNA17, 2'OMe RNA20 and peg-SH (60µl, 700µM), vortexed and left at 25°C for 1 hour mixing at 750rpm. 5nm GNP-BSPP (50µl, 1.74µM) was aliquoted out into three tubes. A solution of reduced 2'OMe RNA ON (Table 4.20) in 0.5 x TBE, 0.015% SDS (300µl) was added to the GNP-BSPP solutions, vortexed and then left at 25°C, 750rpm for 20 minutes.

Table 4.20. RNA ONs RNA18 and RNA19 used in the preparation of GNP16-18 and the volume of reduced 2'OMe RNA ON added to the GNP solution

Reference	Particle Size	Strands Conjugated	Volume of reduced aqueous 2'OMe RNA ON (646µM) added to NP (µl)
GNP16	5nm	RNA17	26.9
GNP17	5nm	RNA17	13.5
		RNA20	13.5
GNP18	5nm	peg-SH	26.9

The reaction was then carried out as in the standard polyconjugation protocol (see Table 4.21, (Section 4.5.1.2) for volumes of salt solution added).

Table 4.21. Volumes of 2M NaCl in 0.5x TBE, 0.015% SDS added to achieve the final concentrations used in splicing assays described in (Section 4.5.1.2)

Salt concentration required	Volume of 2M NaCl in 0.5x TBE, 0.015% SDS (μl)	New total volume (μl)
50mM	9.7	387
100mM	9.9	397
200mM	20.9	417
300mM	22.0	439
400mM	23.1	463
500mM	24.3	486.8
600mM	25.6	512.4
700mM	27.0	539.4

All the particles concentrations were determined and the concentration adjusted so that each solution had a nanoparticle concentration of 1μM for the splicing reactions. Polyconjugated particles were examined by 2% agarose gel at 100V for 1 hour (Section 4.4.3, Figure 4.5).

4.5.4 – GNP Splicing

4.5.4.1 – SMN2 Transcription

A stock transcription reaction was prepared which had six reactions in one, giving transcription volumes of:

8μl Hot transcription buffer

10μl water

4µl Cap

4µl low GTP rNTP mix

4µl *SMN2* mini gene (50ng/µl from PCR)

2µl 0.1M DTT

2µl RNaseOUT TM

2µl T7 polymerase

4µl ³²P GTP

The reaction was carried out as described in the hot transcription protocol (Section 4.5.1.3).

4.5.4.2 – Splicing of 18nm Polyconjugated GNP 200-1nM

NE was dialysed as described in the nanoparticle *in vitro* splicing protocol (Section 4.5.1.4).

Table 4.22. GNP1-9 conjugates tested in a *SMN2* splicing reaction

Reference	Particle Size	Strands conjugated
GNP1	18nm	RNA20
GNP2	18nm	RNA19
GNP3	18nm	RNA18
GNP4	18nm	RNA17
GNP5	18nm	RNA19 RNA20
GNP6	18nm	RNA17 RNA20
GNP7	18nm	RNA18 RNA19
GNP8	18nm	RNA17 RNA18
GNP9	18nm	peg-SH

During the dialysis stocks of 1 μ M, 0.5 μ M, 0.25 μ M, 0.05 μ M and 0.005 μ M of the GNP conjugates GNP1-9 (Table 4.22) were prepared to give final concentrations of nanoparticles in the splicing reaction of 200nM, 100nM, 50nM, 10nM and 1nM respectively. RNA6, GGA O and no oligo were run as controls for this reaction. 1 μ l of these stocks were then treated with 0.5 μ l 32 P labelled RNA and incubated in a 30°C oven for 10min. The reactions were then placed on ice and treated with the master mix as described in the GNP *in vitro* splicing protocol before being placed in a 30°C oven for 2 hrs. After the splicing had finished the reactions were worked up as described in the nanoparticle *in vitro* splicing protocol.

4.5.4.3 – SMN2 Transcription

A transcription reaction was prepared which had six reactions in one, giving transcription volumes of:

8 μ l Hot transcription buffer

10 μ l water

4 μ l Cap

4 μ l low GTP rNTP mix

4 μ l SMN2 mini gene (50ng/ μ l from PCR)

2 μ l 0.1M DTT

2 μ l RNaseOUT™

2 μ l T7 polymerase

4 μ l 32 P GTP

The reaction was carried out as described in the hot transcription protocol (Section 4.5.1.3).

4.5.4.4 – 20nm, 10nm and 5nm GNP10-18 Splicing

NE was dialysed as described in the nanoparticle *in vitro* splicing protocol. During the dialysis stocks of 1µM, 0.5µM, 0.25µM, 0.05µM, 0.005µM and 0.002µM were made up, of the GNP10-18 in Table 4.23, to give final concentrations of nanoparticles in the splicing reaction of 200nM, 100nM, 50nM, 10nM, 1nM and 0.4nM respectively. GGA O, no oligo and dialysed control were the controls for this reaction. 1µl of these stocks were then treated with 0.5µl ³²P labelled RNA and incubated in a 30°C oven for 10min. The reactions were then placed on ice and treated with the master mix as described in the nanoparticles *in vitro* splicing protocol before being placed in a 30°C oven for 2 hours. After the splicing had finished the reactions were worked up as described in the nanoparticle *in vitro* splicing protocol (Section 4.5.1.4).

Table 4.23. GNP10-18 conjugates used in a SMN2 splicing reaction

Reference	Particle Size	Strands conjugated
GNP10	20nm	RNA17
GNP11	20nm	RNA17 RNA20
GNP12	20nm	peg-SH
GNP13	10nm	RNA17
GNP14	10nm	RNA17 RNA20
GNP15	10nm	peg-SH
GNP16	5nm	RNA17
GNP17	5nm	RNA17 RNA20
GNP18	5nm	peg-SH

Chapter 5 – Main finding of the project

5 - Main Findings from this Thesis

5.1 – Discussion

Purification of the 2'OMe RNA1 strands proved problematic as the purification by RP-HPLC (Figure 2.6) could not provide separation due to the presence of an intermolecular parallel quadruplex (Figure 2.32 & Table 2.10). Ion exchange purification however allowed for efficient purification of the 2'OMe RNA1 strands (Figure 2.5), with ion exchange allowing isolation of linear RNA1 strands and only isolation of the folded RNA1 strands. Successful denaturation of the RNA1 strands was achieved by heating at 80°C prior to sample loading onto 20% denaturing polyacrylamide gels (Figure 2.10).

Too low concentrations of CuTBTA stopped the Cu-catalysed Huisgen [3+2] click chemistry reaction from occurring, with 100eq of CuTBTA providing the greatest product conversion. Any attempts to optimise product conversion were unsuccessful with the only method reported in literature to increase product conversion requiring the use of a scaffold strand to facilitate the reaction¹⁴⁹. As in this thesis the optimal separation was being investigated it was more cost effective to use less efficient reactions than to make scaffold strands for every product. Purification of the final clicked RNA strands were carried out by gel electrophoresis, giving yields of ~10%, as this purification technique provided the greatest product recovery.

Splicing experiments carried out with RNA3-RNA14 all increased exon 7 inclusion levels above basal splicing (Table 3.1), indicating that the triazole bioconjugation group in tripartite sequences is tolerated with and without HEG units. Increasing the flexibility of the tripartite strand by addition of 1 or 2 HEG

(RNA6 and RNA7 respectively) gave rise to higher exon 7 inclusion levels than were observed with GGA-O (Table 3.1). However, increasing the flexibility to much decreased activity (Table 3.1). The position of the triazole also exerted an effect on enhancer activity, with moving the triazole from the enhancer side (RNA6) to the annealer side (RNA4), diminishing enhancer activity for exon 7 inclusion (Table 3.1). The proposed rational for the effect seen by modifying the position of the triazole could be that the triazole interfered with protein binding or reduced the looping ability of the enhancer both effecting exon 7 inclusion levels. Maintaining the length of RNA6 but decreasing the flexibility as in RNA14 gave near identical levels of exon 7 inclusion (27% and 26%, respectively) suggesting that length of the RNA strands is involved in the increased exon 7 inclusion levels. However length is not solely the cause, as exchanging the HEG units for abasics, but maintaining the triazole, caused a drop in exon 7 inclusion levels (Table 3.3). Therefore, it seems that both the length and flexibility of RNA strands effects exon 7 inclusion levels in *SMN2* splicing.

The reason phosphothioate backbones (pS) were investigated were because they are reported in literature to increase protein binding². This was further confirmed by experiments carried out in the Eperon lab which found that adding a pS backbone to strands increased activity of TOES strands to exon 7 inclusion levels.

pS Cu-catalysed Huisgen [3+2] click chemistry reactions behaved exactly the same as the non-pS click chemistry reactions, with reactions not going to completion. The secondary structures formed by the pS RNA1 series however were stronger than those formed by the 2'OMe RNA1 sequences, having T_m of 66-64°C and 60-57°C, respectively (Table 2.9). From splicing experiments

carried out it was discovered that full pS backbones had a detrimental effect on splicing activity however strategic placement of pS strands seemed to increase splicing activity^{2,196}. Strategic placements of pS where trialled on RNA16 (a pS version of RNA14) (*Section 3.5.1.6*) increased exon 7 inclusion levels significantly (Table 3.6) showing that pS modification are beneficial for high levels of exon 7 inclusion.

The effectiveness of a multivalent system depends on two main things: particle size and surface coating. With particle sizes of 20nm, 18nm and 10nm, at high concentrations, GNP surface coatings had less of an effect on exon 7 inclusion levels than at lower concentrations. However, at lower concentrations, the RNA17 monolayer coating on GNP4 and mixed monolayer coating of RNA17 and RNA20 on GNP6 did give rise to an increase in exon 7 inclusion levels at GNP concentrations of 100nM and 50nM (*Section 4.4.5.1*, Table 4.4). The directionality of conjugation, like in the tripartite sequence case, also plays a defining role in enhancer activity as shown by GNP2 and GNP4 which have the same RNA sequence, which is conjugated through the 5' end in GNP2 and through the 3' end in GNP4 (*Section 4.4.3*, Table 4.3). The difference the directionality causes in a splicing assay is a shift in exon 7 inclusion from 47% in the presence of GNP4 at 50nM to 34% in the presence of GNP2 at 50nM (*Section 4.4.5.1*, Table 4.4). In the 5nm particles case (GNP16-18) inhibition was still seen at high concentration of GNP however the surface coating exerted an effect by maintaining activity. GNP16, with only an enhancer coating conjugated to the GNP through the 3' end, achieved 80% percentage of exon 7 inclusion at 200nM (*Section 4.4.5.5*, Table 4.6). Other examples of 5nm GNP stimulated exon 7 inclusion to 29% and 39% indicating that with 5nm GNPs, the

coating is essential for enhancer activity and that a higher effective concentration of the enhancer on the surface leads to increased exon 7 inclusion. This indicates that particle size, surface coating and directionality of conjugation are all important for splicing enhancement.

As GNP16 does not have any annealing sequences to direct it to a specific region of pre-mRNA, it was proposed that GNP16 could be used to stimulate splice site inclusion in other systems. GNP16 increased D/S splicing in both Adenovirus mini genes *A2* and *A3*, with the greatest increase being achieved with 200nM GNP16 (*Section 4.4.6*, Table 4.7). Therefore it is concluded that GNP conjugates containing enhancer sequences, in general, act as enhancers for alternative RNA splicing.

5.2 – Conclusion

The tripartite sequences approach is not only a powerful tool for screening enhancer and annealer sequences to be able to rectify aberrant splicing in disease pathways it is also a new method for conjugating 2'OMe RNA and pS RNA strands together. Coupling yields for these conjugation reactions are currently low however can be optimised, using template strands, once the desired length tripartite sequence has been determined for *SMN2* splicing.

All 2'OMe RNA tripartite sequences tested increased exon 7 inclusion levels above basal levels showing that HEG units are tolerated in tripartite sequences. In the case of RNA6 and RNA7, HEG units are actually beneficial in increased exon 7 inclusion levels. Less flexible linkers (abasic linkers) the same length as HEG gave lower levels of exon 7 inclusion, suggesting that the flexibility of the HEG linker plays an important role in the enhancement of exon 7 inclusion.

RNA16 tripartite sequence was even more effective at increasing exon 7 inclusion than all the other sequences due to the strategic placement of pS backbones. The greatest 2'OMe RNA tripartite sequence RNA6 could have potentially increase exon 7 inclusion to higher levels than those obtained with TOES GGA, if a version could have been made with the strategic pS backbone.

Fully phosphothioate versions of tripartite sequences have less activity than non-phosphothioate versions. The reason for this is still under investigation, but may be due to the affect that a pS backbone has on the strength of the binding between annealer sequences and mRNA.

Multivalent system GNP16 is more effective at increasing exon 7 inclusion levels in *SNM2* and Andenovirus D/S splicing levels of *A2* and *A3* than monovalent system TOES. Therefore it is highly probable that GNP16 could be used as a general method for inducing splicing. GNP16, with an RNA17 monolayer coating, was the most effective splicing enhancer of all those produced. GNP16 works effectively as a general enhancer, but attaching annealer sequences onto it, significantly reduced its overall activity. This problem would need to be overcome in order to enable GNP16 to bind specifically to pre-mRNA, without reducing its activity.

5.3 - Future Work

Future work area's to investigate are as follows:-

1. Make a RNA6a versions of RNA6 which has the modified strategically placed pS backbone to investigate how efficient those tripartite sequences would be at including exon 7 in spliced products .

2. Trial click chemistry reactions with template strands to determine whether in this system templates could be used to scale-up production of desired strands.
3. Test the tripartite system in other disease systems using the conjugation group effectively to build large libraries of compounds that able to be conjugated in lab together in 15 minutes.
4. Investigate the quadruplex structure further to determine the type of secondary structure formed and whether the structure is a intermolecular tetrad/hexad type structure as the literature would suggest.
5. Use the enhancer sequence to develop and investigate quadruplex stabilisers. The melting point of the quadruplex is low enough to become a tool for testing the effect quadruplex stabilisers have on the stability of the quadruplex. RNA quadruplexes tend to be really rather stable and having a quadruplex with a lower T_m would help to investigate RNA quadruplex stabilisers
6. Carry out some binding studies on pS strands to determine what effect pS backbones have on RNA-RNA binding. This would prove whether the hypothesis for the weak activity of the fully pS strands is due to binding.
7. Investigate why a mixed monolayer GNP, GNP17, has less activity than the enhancer only strand. During this investigation smaller nanoparticles would be used to see if the low exon 7 inclusion levels observed was due to steric blocking of key sites by the GNP.

Appendix

The Appendix contains:-

- NMR Characterisation data for linkers
- MALDI-tof and accurate mass data for 2'OMe RNA sequences synthesised on the DNA synthesiser
- HPLC traces for RNA2 series purification
- Ion exchange traces for RNA1 series purification
- T_m melting curves for experiments carried out by Alex Cousins
- NMR spectra for RNA sequences carried out by Alex Cousins
- Splicing gels for splicing experiments carried out
- Sequence of SMN2 mini gene
- Sequence of SMN2 transcript

References

- (1) Schneider, M.; Will, C. L.; Anokhina, M.; Tazi, J.; Urlaub, H.; Luhrmann, R. *Molecular Cell* **2010**, 38, 223–235.
- (2) Owen, N.; Zhou, H.; Malygin, A. a; Sangha, J.; Smith, L. D.; Muntoni, F.; Eperon, I. C. *Nucleic Acids Research* **2011**, 39, 7194–208.
- (3) Wahl, M. C.; Will, C. L.; Lu, R. *Cell* **2009**, 136, 701–718.
- (4) Bock, V. D.; Hiemstra, H.; van Maarseveen, J. H. *European Journal of Organic Chemistry* **2006**, 2006, 51–68.
- (5) Ma, M. Y.; Reid, L. S.; Climie, S. C.; Lin, W. C.; Kuperman, R.; Sumner-Smith, M.; Barnett, R. W. *Biochemistry* **1993**, 32, 1751–1758.
- (6) Eritja, R.; Walker, P. A.; Randall, S. K.; Goodman, M. F.; Kaplan, B. E. *Nucleosides, Nucleotides & Nucleic acids* **1987**, 6, 803–814.
- (7) Pennisi, E. *Science* **2003**, 300, 1484.
- (8) Burghes, A. H. M.; Beattie, C. E. *Nature reviews. Neuroscience* **2009**, 10, 597–609.
- (9) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walters, P. *Molecular Biology of the Cell*; Fifth Edit.; 2008.
- (10) Pomeranz Krummel, D. A.; Oubridge, C.; Leung, A. K. W.; Li, J.; Nagai, K. *Nature* **2009**, 458, 475–480.
- (11) Matlin, A. J.; Clark, F.; Smith, C. W. J. *Nature Reviews. Molecular Cell Biology* **2005**, 6, 386–398.
- (12) Zhuang, Y.; Weiner, A. M. *Cell* **1986**, 46, 827–835.
- (13) Seraphin, B.; Rosbash, M. *Cell* **1989**, 59, 349–358.
- (14) Siliciano, P. G.; Guthrie, C. *Genes & Development* **1988**, 2, 1258–1267.
- (15) Kohz, J. D.; Jamison, S. F.; Will, C. L.; Zuo, P.; Luhrmann, R.; Garcia-Blanco, M. A.; Manley, J. L. *Nature* **1994**, 368, 119–124.
- (16) Du, H.; Rosbash, M. *Nature* **2002**, 419, 86–90.
- (17) Patel, S. B.; Bellini, M. *Nucleic Acids Research* **2008**, 36, 6482–6493.
- (18) Raker, V. A.; Hartmuth, K.; Kastner, B.; Luhrmann, R. *Molecular and Cellular Biology* **1999**, 19, 6554–6565.

- (19) Stark, H.; Dube, P.; Luhrmann, R.; Kastner, B. *Nature* **2001**, *409*, 539–542.
- (20) Weber, G.; Trowitzsch, S.; Kastner, B.; Luhrmann, R.; Wahl, M. C. *EMBO Journal* **2010**, *29*, 4172–4184.
- (21) Raker, V. A.; Plessel, G.; Luhrmann, R. *EMBO Journal* **1996**, *15*, 2256–2269.
- (22) Fischer, U.; Sumpter, V.; Sekine, M.; Satoh, T.; Luhrmann, R. *EMBO Journal* **1993**, *12*, 573–583.
- (23) Hamm, J.; Darzynkiewicz, E.; Tahara, S. M.; Mattaj, W. *Cell* **1990**, *62*, 569–577.
- (24) Mattaj I. W.; Boelens, W.; Jarmolowski, A.; Kambach, C. *Molecular Biology Reports* **1993**, *18*, 79–83.
- (25) Branlant, C.; Krol, A.; Ebel, J.; Lazar, E.; Haendler, B.; Jacob, M. *EMBO Journal* **1982**, *1*, 1259–1265.
- (26) Kambach, C.; Walke, S.; Young, R.; Avis, J. M.; Fortelle, E. D.; Raker, V. A.; Lu, R.; Li, J.; Nagai, K. *Cell* **1999**, *96*, 375–387.
- (27) Hamm, J.; Kazmaier, M.; Mattaj, W. *EMBO Journal* **1987**, *6*, 3479–3485.
- (28) Mount, S. M.; Steitz, J. A. *Nucleic Acids Research* **1981**, *9*, 6351–6368.
- (29) Nelissen, R. L. H.; Will, C. L.; Venrooij, W. J. V.; Luhrmann, R. *EMBO journal* **1994**, *13*, 4113–4125.
- (30) Bach, M.; Krol, A.; Lihrmann, R. *Nucleic Acids Research* **1990**, *18*, 449–458.
- (31) Berglund, J. A.; Chua, K.; Abovich, N.; Reed, R.; Rosbash, M. *Cell* **1997**, *89*, 781–787.
- (32) Zhuang, Y.; Goldstein, A. M.; Weinert, A. M. *Genetics* **1989**, *86*, 2752–2756.
- (33) Liu, Z.; Luyten, I.; Bottomley, M. J.; Messias, A. C.; Houngninou-molango, S.; Sprangers, R.; Zanier, K.; Kramer, A.; Sattler, M. *Science* **2001**, *294*, 1098–1102.
- (34) Selenko, P.; Gregorovic, G.; Sprangers, R.; Stier, G.; Rhani, Z.; Kramer, A.; Sattler, M. *Molecular Cell* **2003**, *11*, 965–976.
- (35) Zamore, P. D.; Patton, J. G.; Green, M. R. *Nature* **1992**, *355*, 609–614.
- (36) Banerjee, H.; Rahn, A.; Davis, W.; Singh, R. *RNA* **2003**, *9*, 88–99.

- (37) Banerjee, H.; Rahn, A.; Gawande, B.; Guth, S.; Valcárcel, J.; Singh, R. *RNA* **2004**, *10*, 240–253.
- (38) Berglund, J. A.; Abovich, N.; Rosbash, M. *Genes & Development* **1998**, *12*, 858–867.
- (39) Gupta, A.; Jenkins, J. L.; Kielkopf, C. L. *Journal of Molecular Biology* **2011**, *405*, 1128–1138.
- (40) Kent, O. A.; Reayi, A.; Foong, L.; Chilibeck, K. A.; MacMillan, A. M. *Journal of Biological Chemistry* **2003**, *278*, 50572–50577.
- (41) Valcarcel, J.; Gaur, R. K.; Singh, R.; Green, M. R. *Science* **1996**, *273*, 1706–1709.
- (42) Wu, S.; Romfo, C. M.; Nilsen, T. W.; Green, M. R. *Nature* **1999**, *402*, 832–835.
- (43) Zorio, Diego, A. R.; Blumenthal, T. *Nature* **1999**, *402*, 835–838.
- (44) Merendino, L.; Guth, S.; Bilbao, D.; Martinez, C.; Valcárcel, J. *Nature* **1999**, *402*, 838–841.
- (45) Chusainow, J.; Ajuh, P. M.; Trinkle-mulcahy, L.; Sleeman, J. E.; Ellenberg, J.; Lamond, A. I. *RNA* **2005**, *11*, 1201–1214.
- (46) Rudner, D. Z.; Kanaar, R.; Breger, K. S.; Rio, D. C. *Molecular and Cellular Biology* **1998**, *18*, 1765–1773.
- (47) Abovich, N.; Rosbash, M. *Cell* **1997**, *89*, 403–412.
- (48) Michaud, S.; Reed, R. *Genes & Development* **1991**, *5*, 2534–2546.
- (49) Gozani, O.; Potashkin, J.; Reed, R. *Molecular and Cellular Biology* **1998**, *18*, 4752–4760.
- (50) Fleckner, J.; Zhang, M.; Valcárcel, J.; Green, M. R. *Genes & Development* **1997**, *11*, 1864–1872.
- (51) Kistler, A. L.; Guthrie, C. *Genes & Development* **2001**, *15*, 42–49.
- (52) Brosi, R.; Hauri, H.; Kramer, A. *Journal of Biological Chemistry* **1993**, *268*, 17640–17646.
- (53) Kramer, A.; Utans, U. *EMBO Journal* **1991**, *10*, 1503–1509.
- (54) Staknis, D.; Reed, R. *Molecular and Cellular Biology* **1994**, *14*, 7670–7682.
- (55) Gozani, O.; Feld, R.; Reed, R. *Genes & Development* **1995**, *10*, 233–243.

- (56) Chiara, M. D.; Gozani, O.; Bennett, M.; Champion-Arnaud, P.; Palandjian, L.; Reed, R. *Molecular and cellular biology* **1996**, 16, 3317–26.
- (57) Xu, Y.; Newnham, C. M.; Kameoka, S.; Huang, T.; Konarska, M. M.; Query, C. C. *EMBO Journal* **2004**, 23, 376–385.
- (58) Makarov, E. M.; Makarova, O. V.; Urlaub, H.; Gentzel, M.; Will, C. L.; Wilm, M.; Luhrmann, R. *Science* **2002**, 298, 2205–2208.
- (59) Chiara, M. D.; Palandjian, L.; Kramer, R. F.; Reed, R. *EMBO Journal* **1997**, 16, 4746–4759.
- (60) Bessonov, S.; Anokhina, M.; Krasauskas, A.; Golas, M. M.; Sander, B.; Will, C. L.; Urlaub, H.; Stark, H.; Luhrmann, R. *RNA* **2010**, 16, 2384–2403.
- (61) Lagerbauer, B.; Achsel, T.; Lührmann, R. *Proceedings of the National Academy of Sciences* **1998**, 95, 4188–92.
- (62) Raghunathan, P. L.; Guthrie, C. *Current biology : CB* **1998**, 8, 847–55.
- (63) Staley, J. P.; Guthrie, C. *Molecular Cell* **1999**, 3, 55–64.
- (64) Xiao, X.; Wang, Z.; Jang, M.; Burge, C. B. *Proceedings of the National Academy of Sciences* **2007**, 104, 18583–18588.
- (65) Blencowe, B. J.; Issner, R.; Nickerson, J. A.; Sharp, P. A. *Genes & Development* **1998**, 12, 996–1009.
- (66) Cáceres, J. F.; Kornblihtt, A. R. *Trends in Genetics* **2002**, 18, 186–93.
- (67) Bourgeois, C. F.; Popielarz, M.; Hildwein, G.; Stevenin, J. *Molecular and Cellular Biology* **1999**, 19, 7347–7356.
- (68) Graveley, B. R. *RNA* **2000**, 6, 1197–1211.
- (69) Tian, M.; Maniatis, T. *Genes & Development* **1994**, 8, 1703–1712.
- (70) Tian, M.; Maniatis, T. *Genes & Development* **1994**, 8, 1703–1712.
- (71) Lear, A. L.; Eperon, L. P.; Wheatley, I. M.; Eperon, I. C. *Journal of Molecular Biology* **1990**, 211, 103–115.
- (72) Sorek, R.; Lev-maor, G.; Reznik, M.; Dagan, T.; Belinky, F.; Graur, D.; Ast, G. *Molecular Cell* **2004**, 14, 221–231.
- (73) Zhuang, Y.; Leung, H.; Weiner, A. M. *Molecular and Cellular Biology* **1987**, 7, 3018–3020.
- (74) Yuo, C. Y.; Weiner, A. . M. *Molecular and Cellular Biology* **1989**, 8.

- (75) Roca, X.; Sachidanandam, R.; Krainer, A. R. *RNA* **2005**, *11*, 683–698.
- (76) Freund, M.; Asang, C.; Kammler, S.; Konermann, C.; Krummheuer, J.; Hipp, M.; Meyer, I.; Gierling, W.; Theiss, S.; Preuss, T.; Schindler, D.; Kjems, J.; Schaal, H. *Nucleic Acids Research* **2003**, *31*, 6963–6975.
- (77) Eperon, I. C.; Ireland, D. C.; Smith, R. A.; Mayeda, A.; Krainer, A. R. *The EMBO journal* **1993**, *12*, 3607–3617.
- (78) Wu, J. Y.; Maniatis, T. *Cell* **1993**, *75*, 1061–1070.
- (79) Kohtz, J. D.; Jamison, S. F.; Will, C. L.; Zuo, P.; Luhrmann, R.; Garcia-blanco, M. A.; Manley, J. L. *Nature* **1994**, *368*, 119–124.
- (80) Kent, O. A.; MacMillan, A. M. *Nature Structural Biology* **2002**, *9*, 576–581.
- (81) Das, R.; Zhou, Z.; Reed, R. *Molecular Cell* **2000**, *5*, 779–787.
- (82) Caceres, J. F.; Krainer, A. R. *The EMBO Journal* **1993**, *12*, 4715–4726.
- (83) Eperon, I. C.; Makarova, O. V.; Mayeda, A.; Munroe, S. H.; Caceres, J. F.; Hayward, D. G.; Krainer, A. R. *Molecular and Cellular Biology* **2000**, *20*, 8303–8318.
- (84) Ge, H.; Manley, J. L. *Cell* **1990**, *62*, 25–34.
- (85) Krainer, A. R.; Conway, G. C.; Kozak, D. *Cell* **1990**, *62*, 35–42.
- (86) Wang, J.; Xiao, S.; Manley, J. L. *Genes & Development* **1998**, *12*, 2222–2233.
- (87) Roscigno, R. F.; Garcia-blanco, M. A. *RNA* **1995**, *1*, 692–706.
- (88) Shen, H.; Green, M. R. *Molecular Cell* **2004**, *16*, 363–373.
- (89) Shen, H.; Green, M. R. *Nature Structural & Molecular Biology* **2007**, *14*, 597–603.
- (90) Graveley, B. R.; Hertel, K. J.; Maniatis, T. *RNA* **2001**, *7*, 806–18.
- (91) Henscheid, K. L.; Voelker, R. B.; Berglund, J. A. *Biochemistry* **2008**, *47*, 449–459.
- (92) Tian, M.; Maniatis, T. *Genes & Development* **1994**, *8*, 1703–1712.
- (93) Tian, M.; Maniatis, T. *Cell* **1993**, *74*, 105–114.
- (94) Zamore, P. D.; Green, M. R. *Biochemistry* **1989**, *86*, 9243–9247.
- (95) Zhu, J.; Krainer, A. R. *Genes & Development* **2000**, *14*, 3166–3178.

- (96) Zhu, J.; Mayeda, A.; Krainer, A. R. *Molecular Cell* **2001**, 8, 1351–1361.
- (97) Zuo, P.; Maniatis, T. *Genes & Development* **1996**, 10, 1356–1368.
- (98) Shen, H.; Green, M. R. *Genes & Development* **2006**, 20, 1755–1765.
- (99) Crispino, J. D.; Blencowe, B. J.; Sharp, P. A. *Science* **1994**, 265, 1866–1869.
- (100) Tarn, W.; Steitz, J. A. *Genes & Development* **1994**, 8, 2704–2717.
- (101) Fu, X.; Mayedat, A.; Maniatis, T.; Krainer, A. R. *Proceedings of the National Academy of Sciences* **1992**, 89, 11224–11228.
- (102) Bai, Y.; Lee, D.; Yu, T.; Chasin, L. a *Nucleic acids research* **1999**, 27, 1126–34.
- (103) Zahler, A. M.; Damgaard, C. K.; Kjems, J.; Caputi, M. *Journal of Biological Chemistry* **2004**, 279, 10077–10084.
- (104) Rooke, N.; Markovtsov, V.; Cagavi, E.; Black, D. L. *Molecular and Cellular Biology* **2003**, 23, 1874–1884.
- (105) Kanopka, A.; Muhlemann, O.; Akusjavi, G. *Nature* **1996**, 381, 535–538.
- (106) Simard, M. J.; Chabot, B. *Molecular and Cellular Biology* **2002**, 22, 4001–4010.
- (107) Yu, Y.; Maroney, P.; Denker, J.; Zhang, X.; Dybkov, O.; Luhrmann, R.; Jankowsky, E.; Chasin, L.; Nilsen, T. *Cell* **2008**, 135, 1224–1236.
- (108) Buratti, E.; Baralle, M.; Conti, L. D.; Baralle, D.; Romano, M.; Ayala, Y. M.; Baralle, F. E. *Nucleic Acids Research* **2004**, 32, 4224–4236.
- (109) Chabot, B.; Blanchette, M.; Lapierre, I.; Branche, H. L. *Molecular and Cellular Biology* **1997**, 17, 1776–1786.
- (110) Shen, J.; Zu, K.; Cass, C. L.; Beyer, A. L.; Hirsh, J. *Biochemistry* **1995**, 92, 1822–1825.
- (111) Yang, X.; Bani, M.; Lu, S.; Rowan, S.; Ben-david, Y.; Chabot, B. *Proc. Natl. Acad. Sci. USA* **1994**, 91, 6924–6928.
- (112) Jacquenet, S.; Mereau, A.; Bilodeau, P. S.; Damier, L.; Stoltzfus, C. M.; Branlant, C. *Journal of Biological Chemistry* **2001**, 276, 40464–40475.
- (113) House, A. E.; Lynch, K. W. *Journal of Biological Chemistry* **2008**, 283, 1217–1221.

- (114) Bonnal, S.; Martinez, C.; Forch, P.; Bachi, A.; Wilm, M.; Valcárcel, J. *Molecular Cell* **2008**, 32, 81–95.
- (115) Sharma, S.; Kohlstaedt, L. A.; Damianov, A.; Rio, D. C.; Black, D. L. *Nature Structural & Molecular Biology* **2008**, 15, 183–191.
- (116) Allain, H.; Black, D. L.; Sharma, S.; Maris, C. *Molecular Cell* **2011**, 41, 579–588.
- (117) Fiset, J.; Toutant, J.; Dugré-brisson, S.; Desgroseillers, L.; Chabot, B. *RNA* **2010**, 16, 228–238.
- (118) Caputi, M.; Zahler, A. M. *EMBO Journal* **2002**, 21, 845–855.
- (119) Cunningham, S. A.; Else, A. J.; Potter, B. V. L.; Eperon, I. C. *Journal of Molecular Biology* **1991**, 217, 265–281.
- (120) Pasman, Z.; Garcia-Blanco, M. a *Nucleic Acids Research* **1996**, 24, 1638–1645.
- (121) Reed, R.; Maniatis, T. *Cell* **1986**, 46, 681–690.
- (122) Reed, R. *Genes & Development* **1989**, 3, 2113–2123.
- (123) Guth, S.; Tange, T. O.; Kellenberger, E.; Valcarcel, J. *Molecular and Cellular Biology* **2001**, 21, 7673–7681.
- (124) Kan, J. L.; Green, M. R. *Genes & Development* **1999**, 13, 462–71.
- (125) Robberson, B. L.; Cote, G. J.; Berget, S. M. *Molecular and Cellular Biology* **1990**, 10, 84–94.
- (126) Fox-walsh, K. L.; Dou, Y.; Lam, B. J.; Hung, S.; Baldi, P. F.; Hertel, K. J. *Proceedings of the National Academy of Sciences* **2005**, 102, 16176–16181.
- (127) McManus, C. J.; Graveley, B. R. *Current opinion in genetics & development* **2011**, 21, 373–9.
- (128) Faustino, N. A.; Cooper, T. A. *Genes & Development*. **2003**, 17, 419–37.
- (129) Lopez-Bigas, N.; Audit, B.; Ouzounis, C.; Parra, G.; Guigó, R. *FEBS Letters* **2005**, 579, 1900–1903.
- (130) Gubitz, A. K.; Feng, W.; Dreyfuss, G. *Experimental Cell Research* **2004**, 296, 51–56.
- (131) Hua, Y.; Vickers, T. a; Baker, B. F.; Bennett, C. F.; Krainer, A. R. *PLoS Biology* **2007**, 5, e73.

- (132) Zhang, Z.; Lotti, F.; Dittmar, K.; Younis, I.; Wan, L.; Kasim, M.; Dreyfuss, G. *Cell* **2008**, *133*, 585–600.
- (133) Skordis, L. a; Dunckley, M. G.; Yue, B.; Eperon, I. C.; Muntoni, F. *Proceedings of the National Academy of Sciences* **2003**, *100*, 4114–4119.
- (134) Cartegni, L.; Hastings, M. L.; Calarco, J. a; de Stanchina, E.; Krainer, A. R. *American Journal of Human Genetics* **2006**, *78*, 63–77.
- (135) Cartegni, L.; Krainer, A. R. *Nature genetics* **2002**, *30*, 377–84.
- (136) Kashima, T.; Rao, N.; Manley, J. L. *Proceedings of the National Academy of Sciences* **2007**, *104*, 3426–3431.
- (137) Prior, T. W.; Krainer, A. R.; Hua, Y.; Swoboda, K. J.; Snyder, P. C.; Bridgeman, S. J.; Burghes, A. H. M.; Kissel, J. T. *American Journal of Human Genetics* **2009**, *85*, 408–13.
- (138) Hastings, M. L.; Berniac, J.; Liu, Y. H.; Abato, P.; Jodelka, F. M.; Barthel, L.; Kumar, S.; Dudley, C.; Nelson, M.; Larson, K.; Bowser, T.; Draper, M.; Higgins, P.; Krainer, A. R. **2010**, *1*, 1–21.
- (139) Skordis, L. a; Dunckley, M. G.; Yue, B.; Eperon, I. C.; Muntoni, F. *Proceedings of the National Academy of Sciences* **2003**, *100*, 4114–9.
- (140) Liu, H.-X.; Zhang, M.; Krainer, a. R. *Genes & Development* **1998**, *12*, 1998–2012.
- (141) Pasman, Z.; Garcia-Blanco, M. a *Nucleic Acids Research* **1996**, *24*, 1638–45.
- (142) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B. *Journal of the American Chemical Society* **2003**, *125*, 3192–3193.
- (143) Gartner, Z. J.; Grubina, R.; Calderone, C. T.; Liu, D. R. *Angewandte Chemie (International ed. in English)* **2003**, *42*, 1370–5.
- (144) Gartner, Z. J.; Liu, D. R. *Journal of the American Chemical Society* **2001**, *123*, 6961–3.
- (145) Gramlich, P. M. E.; Warncke, S.; Gierlich, J.; Carell, T. *Angewandte Chemie (International ed. in English)* **2008**, 3442–3444.
- (146) Liu, P.-Y.; Jiang, N.; Zhang, J.; Wei, X.; Lin, H.-H.; Yu, X.-Q. *Chemistry & Biodiversity* **2006**, *3*, 958–66.
- (147) Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G. *Angewandte Chemie (International ed. in English)* **2009**, *48*, 9879–83.

- (148) Kumar, R.; El-Sagheer, A.; Tumpene, J.; Lincoln, P.; Wilhelmsson, L. M.; Brown, T. *Journal of the American Chemical Society* **2007**, 129, 6859–64.
- (149) El-Sagheer, A. H.; Brown, T. *Proceedings of the National Academy of Sciences* **2010**, 107, 15329–34.
- (150) El-Sagheer, A. H.; Brown, T. *Chemical Society Reviews* **2010**, 39, 1388–405.
- (151) Wu, M.; Tinoco, I. *Proceedings of the National Academy of Sciences* **1998**, 95, 11555–60.
- (152) Joachimi, A.; Benz, A.; Hartig, J. S. *Bioorganic & medicinal chemistry* **2009**, 17, 6811–5.
- (153) Xu, Y.; Suzuki, Y.; Komiyama, M. *Angewandte Chemie (International ed. in English)* **2009**, 48, 3281–4.
- (154) Lipay, J. M.; Mihailescu, M.-R. *Molecular BioSystems* **2009**, 5, 1347–55.
- (155) Mergny, J.-L.; De Cian, A.; Ghelab, A.; Saccà, B.; Lacroix, L. *Nucleic Acids Research* **2005**, 33, 81–94.
- (156) Palumbo, S. L.; Memmott, R. M.; Uribe, D. J.; Krotova-khan, Y.; Hurley, L. H.; Ebbinghaus, S. W. *Nucleic Acids Research* **2008**, 36, 1755–1769.
- (157) Krotz, A.; Hang, A.; Gorman, D.; Scozzari, A. *Nucleosides, Nucleotides & Nucleic Acids* **2005**, 24, 1293–1299.
- (158) Weiss, P. A.; Jenny, L.; Stutz, A.; Wu, X. *Helvetica Chimica Acta* **2001**, 84, 3773–3795.
- (159) Tung, C. H.; Stein, S. *Bioconjugate Chemistry* **2000**, 11, 605–618.
- (160) Krotz, A. H.; Gorman, D.; Mataruse, P.; Foster, C.; Godbout, J. D.; Coffin, C. C.; Scozzari, A. N. *Organic Process Research & Development* **2004**, 8, 852–858.
- (161) Vacek, M.; Sazani, P.; Kole, R. *Cellular and Molecular Life Sciences : CMLS* **2003**, 60, 825–33.
- (162) Scaffidi, P.; Misteli, T. *Nature Medicine* **2005**, 11, 440–5.
- (163) Wilton, S. D.; Fletcher, S. *Current Gene Therapy* **2005**, 5, 467–486.
- (164) van Deutekom, J. C. T.; van Ommen, G.-J. B. *Nature Reviews. Genetics* **2003**, 4, 774–83.

- (165) Lu, Q. L.; Rabinowitz, A.; Chen, Y. C.; Yokota, T.; Yin, H.; Alter, J.; Jadoon, A.; Bou-Gharios, G.; Partridge, T. *Proceedings of the National Academy of Sciences* **2005**, *102*, 198–203.
- (166) Alter, J.; Lou, F.; Rabinowitz, A.; Yin, H.; Rosenfeld, J.; Wilton, S. D.; Partridge, T. a; Lu, Q. L. *Nature Medicine* **2006**, *12*, 175–7.
- (167) Singh, N. K.; Singh, N. N.; Androphy, E. J.; Singh, R. N. *Molecular and Cellular Biology* **2006**, *26*, 1333–1346.
- (168) Singh, N. N.; Androphy, E. J.; Singh, R. N. *Biochemical and Biophysical Research Communications* **2004**, *315*, 381–8.
- (169) Miyajima, H.; Miyaso, H.; Okumura, M.; Kurisu, J.; Imaizumi, K. *Journal of Biological Chemistry* **2002**, *277*, 23271–7.
- (170) Madocsai, C.; Lim, S. R.; Geib, T.; Lam, B. J.; Hertel, K. J. *Molecular Therapy* **2005**, *12*, 1013–22.
- (171) Cartegni, L.; Krainer, A. R. *Nature Structural Biology* **2003**, *10*, 120–5.
- (172) Barash, Y.; Calarco, J. A.; Gao, W.; Pan, Q.; Wang, X.; Shai, O.; Blencowe, B. J.; Frey, B. J. *Nature* **2010**, *465*, 53–59.
- (173) Massich, M. D.; Giljohann, D. A.; Schmucker, A. L.; Patel, P. C.; Mirkin, C. A. *ACS NANO* **2010**, *4*, 5641–5646.
- (174) Giljohann, D. A.; Seferos, D. S.; Prigodich, A. E.; Patel, P. C.; Mirkin, C. A. *Journal of the American Chemical Society* **2009**, *131*, 2072–2073.
- (175) Giljohann, D. A.; Seferos, D. S.; Daniel, W. L.; Massich, M. D.; Patel, P. C.; Mirkin, C. A. *Angewandte Chemie (International ed. in English)* **2010**, *49*, 3280–94.
- (176) Prigodich, A. E.; Seferos, D. S.; Massich, M. D.; Giljohann, D. A.; Lane, B. C.; Mirkin, C. A. *ACS NANO* **2009**, *3*, 2147–2152.
- (177) Patel, P. C.; Giljohann, D. A.; Seferos, D. S.; Mirkin, C. A. *Proceedings of the National Academy of Sciences* **2008**, *105*, 17222–17226.
- (178) Larginho, M.; Baptista, P. V. *Journal of Proteomics* **2011**, *75*, 2811–2823.
- (179) Ryou, S.; Kim, J.; Yeom, J.; Hyun, S.; Kim, S.; Su, M.; Wouk, S.; Bae, J.; Rhee, S.; Lee, K. *Biochemical and Biophysical Research Communications* **2011**, *416*, 178–183.
- (180) Kim, D.; Kim, J.; Park, M.; Yeom, J.; Go, H.; Kim, S.; Su, M.; Lee, K.; Bae, J. *Biomaterials* **2011**, *32*, 2593–2604.

- (181) Liu, Y.; Franzen, S. *Bioconjugate Chemistry* **2008**, *19*, 1009–16.
- (182) Giljohann, D. A.; Seferos, D. S.; Patel, P. C.; Millstone, J. E.; Rosi, N. L.; Mirkin, C. A. *Nano Letters* **2007**, *7*, 3818–3821.
- (183) Seferos, D. S.; Prigodich, A. E.; Giljohann, D. A.; Patel, P. C.; Mirkin, C. A. *Nano Letters* **2009**, *9*, 308–311.
- (184) Rosi, N. L.; Mirkin, C. A. *Chem. Rev* **2005**, *105*, 1547–1562.
- (185) Seferos, D. S.; Giljohann, D. A.; Hill, H. D.; Prigodich, A. E.; Mirkin, C. A. *Journal of the American Chemical Society* **2007**, *129*, 15477–15479.
- (186) Zheng, D.; Seferos, D. S.; Giljohann, D. A.; Patel, P. C.; Mirkin, C. A. *Nano Letters* **2009**, *9*, 3258–3261.
- (187) Hurst, S. J.; Lytton-Jean, A. K. R.; Mirkin, C. A. *Analytical Chemistry* **2006**, *78*, 8313–8318.
- (188) Miele, E.; Spinelli, G. P.; Miele, E.; Di Fabrizio, E.; Ferretti, E.; Tomao, S.; Gulino, A. *International Journal of Nanomedicine* **2012**, *7*, 3637–57.
- (189) Kumar, S.; Aaron, J.; Sokolov, K. *Nature Protocols* **2008**, *3*, 314–320.
- (190) Zhao, W.; Lin, L.; Hsing, I. *Bioconjugate chemistry* **2009**, *20*, 1218–1222.
- (191) Cioran, A. M.; Musteti, A. D.; Teixidor, F.; Krpetić, Ž.; Prior, I. a; He, Q.; Kiely, C. J.; Brust, M.; Viñas, C. *Journal of the American Chemical Society* **2012**, *134*, 212–21.
- (192) Krpetić, Ž.; Nativo, P.; Prior, I. A.; Brust, M. *Small* **2011**, *7*, 1982–1986.
- (193) Dougan, J. a; Karlsson, C.; Smith, W. E.; Graham, D. *Nucleic Acids Research* **2007**, *35*, 3668–75.
- (194) Technologies, L. *Technical Information for Products*; 2010; pp. 1–144.
- (195) BBI International *GNP extinction coefficients*; 2004; p. 1.
- (196) Perrett, A. J.; Dickinson, R. L.; Krpetić, Ž.; Brust, M.; Lewis, H.; Eperon, I. C.; Burley, G. a. *Chemical Science* **2013**, *4*, 257.