Modification of splicing with antisense oligonucleotides in the Insulin receptor exon 11 and Apolipoprotein B exon 26.

Srirangalingam, Umasuthan

Modification of splicing with antisense oligonucleotides in the Insulin receptor exon 11 and Apolipoprotein B exon 26.

Umasuthan Srirangalingam

BSc. (Hons), MB ChB, MRCP

A thesis submitted to the University of London for the degree of Doctor of Philosophy

Centre for Endocrinology,
St Bartholomew's Hospital,
Queen Mary and Westfield University of London

Declarations

I declare that the work presented in this thesis is original work carried out by Umasuthan Srirangalingam in the Centre for Endocrinology, St Bartholomew's Hospital and Queen Mary, University of London.

Umasuthan Srirangalingam.

Abstract

Background - The alternatively spliced insulin receptor (IR) exon 11 (36 nucleotides) and the constitutively spliced Apolipoprotein B (APOB) exon 26 (7572 nucleotides) are examples of the shortest and longest exons in the genome.

Aim - The aim of this study was to investigate the regulation of splicing of these 2 exons in cell culture using 2′-O-methyl RNA antisense oligonucleotides (ASOs) and peptide nucleic acid (PNA)-peptide hybrid ASOs.

Methods - ASOs were designed to target key sequences involved in the splicing of the IR exon 11 and exonic splicing silencer sequences (ESS) in APOB exon 26. HepG2 cells were reverse-transfected with the ASOs for 48 hours, mRNA harvested and RT-PCR was performed to amplify the IR isoform and APOB cDNAs which were separated by PAGE and quantified.

Results

Insulin receptor exon 11 - 2′-O-methyl RNA ASOs targeted to two intronic sites, the 3′ half of exon 11 and spanning the entire exon caused significant exon skipping.

PNA-peptide hybrids predicted to increase exon 11 splicing, paradoxically caused exon skipping. PNA-peptide hybrids with 3′ tails caused exon 11 skipping more effectively than hybrids with 5′ tails.

Apolipoprotein B exon 26 - Only combinations of 2′-O-methyl RNA ASOs targeting multiple ESSs in APOB exon 26 caused a small proportion of aberrant

splicing. This consisted of complete exon 26 skipping and the selection of a downstream cryptic 3′ splice site in preference to the native 3′ splice site.

Discussion - Exclusion of the IR exon 11 can be induced by targeting a combination of intronic or exonic sequences. PNA-peptide hybrid ASOs were unable to increase exon 11 splicing. The aberrant splicing of large constitutive exons such as APOB exon 26 can be induced by targeting multiple ESS sites along its course.

Acknowledgements

I am grateful to my supervisors, Professor Shern Chew and Dr Bernard Khoo for their, patience, commitment, guidance and ongoing support throughout my research.

I would like to acknowledge the support of both the Joan Adams Fellowship and the Barts and the London Charity for grant support to carry out this research.

Finally, I would like to thank my mother and father, my wife Radha and of course little Ambihai and Ramesh for their love and support during this time.

Table of Contents

List of figures

3.1.4.C Graphical representation: Fluorescein-tagged 2′-O-methyl RNA ASOs targeting the IR exon 11.

3.1.4.D Confocal microscopy: Fluorescein-tagged 2′-O-methyl RNA ASOs targeting the IR exon 11.

3.1.5.A Schematic figure: TOES ASOs targeting the IR exon 11

3.1.5.B Polyacrylamide gel: TOES ASOs targeting the IR exon 11

3.1.5.C Graphical representation: TOES ASOs targeting the IR exon 11

3.2.1.A Schematic figure: PNA ASOs targeting the IR exon 11

3.2.1.B Polyacrylamide gel: PNA ASOs targeting the IR exon 11

3.2.1.C Graphical representation: PNA ASOs targeting the IR exon 11

3.2.2.A Schematic figure: PNA ASOs targeting an equivalent sequence to ASO INSR ex2 in the IR exon 11.

3.2.2.B Polyacrylamide gel: PNA ASOs targeting an equivalent sequence to ASO INSR ex2 in the IR exon 11.

3.2.2.C Graphical representation: PNA ASOs targeting an equivalent sequence to ASO INSR ex2 in the IR exon 11.

3.2.3.A Schematic figure: PNA-peptide hybrid ASOs (5′ tail) targeting the 5′ end of IR exon 11.

3.2.3.B Polyacrylamide gel: PNA-peptide hybrid ASOs (5′ tail) targeting the 5′ end of IR exon 11.

3.2.3.C Graphical representation: PNA-peptide hybrid ASOs (5′ tail) targeting the 5′ end of IR exon 11.

3.2.3.D Polyacrylamide gel: PNA-peptide hybrid ASOs (3′ tail) targeting the 5′ end of IR exon 11.

3.2.3.E Graphical representation: PNA-peptide hybrid ASOs (3′ tail) targeting the 5′ end of IR exon 11.

Chapter 4.0 Discussion

4.2.1.A Schematic figure: Proposed model for interaction between IR exon 11 and splicing factors.

4.2.1.B Schematic figure: Splicing factors and ASOs - interaction with the IR exon 11

List of tables

Abbreviations

AEC	A-like exon definition complex
AIDS	Acquired immunodeficiency syndrome
ANOVA	Analysis of variance
APE1	Apurinic⁄apuridinic endonuclease 1
APOB	Apolipoprotein B
APOBEC 1	Apolipoprotein B mRNA editing enzyme catalytic polypeptide 1
APOE	Apolipoprotein E
APS	Ammonium persulphate
ARD-1	Activator of RNA decay
ASO	Antisense oligonucleotides
ATI	Alternative transcription initiation
ATP	Adenosine-5′ -triphosphate
ATT	Alternative transcription termination
Bcl	B-cell lymphoma
BPS	Branch point sequences
BRCA 1	Breast cancer 1 gene
cDNA	complementary DNA
CMV	Cytomegalovirus
CoAA	Coactivator activator
CoSMoS	Co-localisation single-molecule spectroscopy
CPSF-73	Cleavage and polyadenylation specificity factor -73kDa
CTD	C terminal domain
CUG-BP1	CUG triplet repeat, RNA binding protein 1,
DAKO	DakoCytomation Fluorescent Mounting Medium

DAPI	4', 6-diamidino-2-phenylindole satin
dH_20	Deionised water
dNTP	Deoxynucleotide triphosphate
DMSO	Dimethyl sulfoxide
DMPK	Dystrophia myotonic protein kinase gene
DNA	Deoxyribonucleic acid
EJC	Exon junction complex
ERK	extracellular-signal-regulated kinase
ESE	Exonic splicing enhancer
ESS	Exonic splicing silencer
ESSENCE	Exon specific splicing enhancement by small chimeric effectors
FBS	Fetal bovine serum
FGFR2	Fibroblast growth factor receptor 2
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GFP	Green fluorescent protein
GLUT4	Glucose transporter type 4
HEK293	Human embryonic kidney cell 293
Hep G2	Human hepatoblastoma cell line G2
hnRNP	Heterogeneous nuclear ribonucleoprotein
HPRT	Hypoxanthine-guanine phosphoribosyltransferase gene
IDL	Intermediate density lipoproteins
IE2	Major immediate early region 2
IGF-I	Insulin-like growth factor I
IGF-II	Insulin-like growth factor II
IL	Intron large

IL-1	interleukin-1
InsP6	Inositol hexakisphosphate
INSR	Insulin receptor
IR	Insulin receptor
IR-A	Insulin receptor - A isoform
IR-B	Insulin receptor - B isoform
IRE1	Inositol requiring enzyme-1
IRS	Insulin receptor substrate
ISE	Intronic splicing enhancer
ISS	Intronic splicing silencer
IVS1	First intervening sequence of the beta-globin gene
Jmjd6	Jumonji domain-6 protein
LDL	Low density lipoprotein
LINE	Long interspersed element repeats
LNA	Locked nucleic acid
LSm	Sm-like protein
MEM	Modified Eagle's medium
MBNL1	Muscleblind-like 1
MCAD	Medium-chain acylCoA dehydrogenase
$MgCl_2$	Magnesium chloride
miRNA	microRNA
mRNA	Messenger ribonucleic acid
mTOR	Mammalian target of rapamycin
NaOAc	Sodium Acetate
NaOH	Sodium Hydroxide
NAS	Nonsense-associated altered splicing

NAT	Natural antisense transcripts
NCAM	Neural cell adhesion molecule gene
NMD	Nonsense mediate decay
NMT1	N-myristoyltransferase 1
NPC	Nuclear pore complex
NTC	NineTeen Complex
PAGE	Polyacrylamide gel
PBS	Phosphate buffered saline
PCR	Polmerase chain reaction
PDK	3′ -phosphoinositide-dependent kinase-1
PIP2	Phosphatidylinositol 4,5-bisphosphate
PIP3	Phosphatidylinositol (3,4,5)-trisphosphate
PKB	Protein Kinase B
PKC	Protein Kinase C
PNA	Peptide nucleic acid
PMO	Phosphorodiamidate morpholino oligonucleotide
PPIases	Peptidyl-prolyl isomerases
PPT	Polypyrimidine tract
PS	Phosphorothioate oligonucleotide
PTC	Premature termination codon
PTB	Polypyrimidine binding tract protein
RBM5	RNA-binding motif protein 5
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNA Pol II	RNA polymerases II
RRM	RNA-recognition motifs

RS	Arginine/serine
RT	Reverse transcription
SELEX	Systematic evolution of ligands by exponential enrichment
SEM	Standard error of mean
SF1	Splicing factor 1
SF2/ASF	Splicing factor 2/ alternative splicing factor
SINE	Short interspersed element repeats
SiRNA	Small interfering RNA
SKAR	S6 kinase 1 Aly/REF-like target
snRNA	Small nuclear RNA
snRNP	Small nuclear ribonucleoprotein
SMA	Spinal muscular atrophy
SMaRT	Spliceosome-mediated RNA trans-splicing
SXL	Sex-lethal
SMN	Survival of motor neurone
TAE	Tris base/acetic acid/EDTA buffer
TBE	Tris/Borate/EDTA buffer
TE	Tris/EDTA solution
TEMED	Tetramethylethylenediamine
Top1	DNA topoisomerase I
TOES	Tailed oligonucleotide enhancement of splicing
TREX	Transcription/export complex
tRNA	Transfer RNA
TPP	Thiamine pyrophosphate
U2AF	U2 Auxiliary Factor
UTR	Untranslated Region

VLDL	Very low density lipoprotein
2'-MOE	2'-*O*-methoxyethyl
2'-OMe	2'-*O*-methyl

Chapter 1

Introduction

1.1 Introduction.

The elucidation of the human genome has had a significant impact on endocrinology. The realisation that the number of human genes is more limited than initially predicted has emphasised the key role of molecular mechanisms that lead to phenotypic diversity, of which RNA splicing is one. Identification of the nucleotide sequences of endocrine genes has allowed an understanding and modulation of their functions. Antisense strategies have developed directly from this knowledge.

This work focuses on understanding gene expression in two key genes in the endocrine system: the Insulin Receptor (IR) and Apolipoprotein B (APOB) genes. Both are involved in key metabolic pathways. This thesis attempts to understand the regulation of the splicing of these two genes to modulate gene expression for possible future therapeutic manoeuvres. This work reports the modification of the splicing process using antisense oligonucleotides (ASOs) in these two contrasting endocrine models of the insulin receptor exon 11 and the Apolipoprotein B exon 26.

1.2 Splicing.

1.2.1 The 'Central Dogma' of Molecular Biology.

The generation of protein from DNA involves the process of transcription (DNA into RNA) and translation (RNA into protein). DNA is initially transcribed by RNA polymerases (RNA Pol II) into pre-messenger RNA (pre-mRNA), which contains both coding exons and non-coding introns. The pre-mRNA is processed at both ends of the transcript with 5' capping as it emerges from RNA Pol II and 3' polyadenylation. Pre-mRNA is spliced (section 1.2.2) to generate the mature messenger RNA (mRNA). The mRNA is exported out of the nucleus into the cytoplasm as part of a messenger ribonucleoprotein (mRNP) and then translat2ed

into an amino acid sequence by the cytoplasmic ribosomes in association with tRNA, to synthesise proteins. The transcript remains under surveillance throughout the process and incomplete or miscopied transcripts are identified and targeted for degradation.

1.2.2 Splicing and alternative splicing.

Splicing is the process by which introns are removed and exons ligated from pre-mRNA sequences to form a translatable message, messenger RNA (mRNA). Alternative splicing is the process by which multiple protein isoforms are generated from one pre-mRNA due to inclusion of alternative exons. Alternative splicing events include exon skipping, intron retention, the use of mutually exclusive exons and alternative 5' and 3' splice sites. The splicing mechanism is reliant on the co-ordinated interaction of both cis (sequences within the pre-mRNA) and trans-acting (molecules and complexes separate from the pre-mRNA) elements.

1.2.3 The need for genetic diversity.

The Human genome consists of 3 billion nucleotide base-pairs making up approximately 22,000 genes, though only 1.2% of this constitutes exonic coding sequence (International Human Genome Sequencing Consortium 2004). The number of genes in the human genome is only twice the number found in a simple organism such as the fruit fly (~14,000 genes) despite the vast difference in organism complexity (Halligan et al. 2006). Mechanisms to account for the diversity of the resultant proteome include the regulation of transcription, splicing, polyadenylation and translation.

Alternative splicing provides a major mechanism for generating phenotypic diversity from such a limited number of genes. Alternative splicing occurs in 95% of human genes in comparison to significantly lower rates in simple organisms and the number of splice variants per gene correlates with the complexity of the organism suggesting an evolutionary drive to efficiently code for complexity within the genome (Wang et al. 2008; Pan et al. 2008; Kim et al. 2004). Comparative genomic studies reinforce this assertion finding that alternatively spliced exons tend not to be conserved between species, suggesting they are a more recent evolutionary development (Modrek et al. 2003).

Alternative transcription includes alternative transcription initiation (ATI) in which an alternative promoter in the 5' untranslated region (UTR) is utilised and alternative transcription termination (ATT) in which an alternative termination site is used. Transcriptome analysis using next generation RNA sequencing (RNA seq) has made it feasible to compare alternative transcriptional to alternative splicing events. A study of the transcriptome of 15 human tissues showed that alternative splicing events predominated over alternative transcriptional events (ATI and ATT) (Wang et al. 2008). However more recent analyses of the transcriptome suggests that alternative transcriptional events predominated over splicing events overall, to account for transcriptome diversity (Pal et al. 2011; Shabalina et al. 2010; Gonzalez-Porta et al. 2012). Possible explanations for this discrepancy include the different species studied (human and mouse), the improved definition of 5' and 3' splice sites and the variable use of multiple genome databases.

The relative frequency of alternative transcription and splicing events varies depending on the location within a transcript, thus alternative splicing is the predominant alternative event in the coding region of the transcript but ATI and ATT

appear to more important overall and specifically at the 5'UTR and 3'UTR respectively (Shabalina et al. 2010).

Alternative translation occurs when translation is initiated by the ribosome at an alternative start codon and can increase genetic diversity by generating different truncated proteins (Kochetov 2008). Recent estimates that up to 10% of human genes have alternative translational initiation sites (Zhang et al. 2007) suggests that this mechanism plays a less prominent role than alternative splicing or transcription.

Gene duplication during DNA replication in mitosis or meiosis is a mechanism that has been postulated to play an important role in contributing to genetic diversity. A study comparing the genomes of nine eukaryotic species concluded that duplicate genes arise at a high rate of 0.01 of a gene per million years (Lynch et al. 2000). Mutations and natural selection may then generate new functional gene products (neofunctionalisation). However the majority of gene duplications do not lead to the generation of new functional proteins. Most will degenerate and be silenced (nonfunctionalisation) with others dividing functions between the duplicated genes (subfunctionalisation). If gene duplication were a dominant mechanism, complex organism may be expected to contain proportionately more genes than simple organisms but this is not the case. Thus while gene duplication remains an important mechanism for generating genetic diversity for species over vast spans of time it plays a negligible role in the variability of gene expression for the individual organism.

1.2.4 The splicing mechanism.

1.2.4.1 Cis-acting sites involved in splicing.

Splicing is determined by the identification of degenerate RNA sequences defining intron/exon boundaries which include the 3′ splice site and 5′ splice site (figure 1.2.4.A). These sequences are recognised, in the majority of cases, by the intronic residues GU at the 5′ splice site and AG at the 3′ splice site in the predominant canonical form of splicing. Non-canonical splicing via the minor spliceosome recognises other sequences at the 5′ and 3′ end of the intron e.g. AT-AC respectively for U12 type introns.

The inherent strength of the 5’ splice site has been estimated based on the sequence of 9 nucleotides at positions -3 to +6 (position 0 defined as the exon-intron boundary) with further studies demonstrating the additional utility of including nucleotides in position 7+ and 8+ (Seraphin et al. 1988; Freund et al. 2003). Bioinformatic tools which identify splice site sequences may do so by assessing the nucleotides in isolation, with respect to the adjacent nucleotide or non-adjacent nucleotides. The position-specific weight matrix of Shapiro and Senapathy gives a score for the strength of a splice site assuming independence between adjacent nucleotide sites (Shapiro et al. 1987). Burge and Karlin produced a predictive tool, GENSCAN, which identifies potential splice sites based on the interdependencies between nucleotides both adjacent and at a distance from one another (Burge et al. 1997).

Other key intronic sites upstream of the 3′ splice site include the polypyrimidine tract (PPT), a sequence rich in the pyrimidines uracil and cytosine. It functions as the site of binding for the splicing factor U2AF65 (Zamore et al. 1989). The BPS is situated further upstream and contains a conserved adenine residue within the consensus sequence YNCUGAC (where N is any nucleotide and Y is a pyrimidine)

(Reed et al. 1988). The BPS binds Splicing factor 1 (SF1) during the formation of the E complex and is necessary for ligation of the intron to form the lariat during splicing (Berglund et al. 1997).

Figure 1.2.4.A

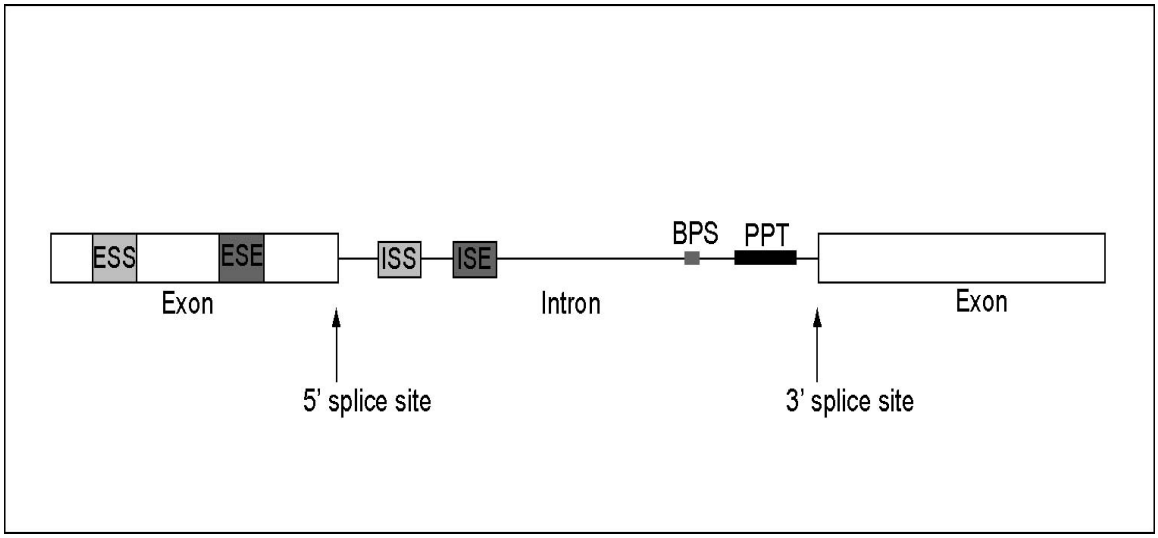


Figure 1.2.4.A Cis-acting splicing elements. Figure shows section of pre-mRNA. Introns are bound by the 5′ splice site and the 3′ splice site. Further key intronic sites involved in the process of splicing include the polypyrimidine tract (PPT) and the branch point sequence (BPS). Regulatory elements in the intron include intronic splicing enhancers (ISE) and intronic splicing silencers (ISS). Exonic regulatory elements include exonic splicing enhancers (ESE) and exonic splicing silencers (ESS).

1.2.4.2 The spliceosome.

The major spliceosome is a nuclear complex which includes both RNA and protein elements in the form of five small nuclear ribonucleoproteins (snRNPs), designated U1, U2, U4, U5 and U6 and many non-snRNPs (Jurica et al. 2003). Over 300 proteins have been shown to co-purify with the splicing complex (Rappsilber et al. 2002). The minor spliceosome is a separate ribonucleoprotein complex involved in the removal of the less common U12 type introns. Both the major and minor spliceosome reside and function within the nucleus of the cell (Pessa et al. 2008). They co-ordinate the identification of intron/exon boundaries within the pre-mRNA and interactions between splicing factors and the pre-mRNA in the subsequent

splicing process. SnRNPs, via their RNA and protein components, are able to mediate RNA-RNA interactions (Watson and Crick base-pairing), RNA-protein and protein-protein interactions within the spliceosome. Proteomic analysis of the spliceosome has identified a significant number of proteins that have known roles in processes other than splicing, emphasising the coupling of these different aspects of gene expression (Zhou et al. 2002).

Detailed analysis of the crystalline structure of U1 snRNP provides further insights to the composition and structure of snRNPs. The U1 snRNP is composed of the 4 stem loops of U1 small nuclear RNA (snRNA) and 7 Sm proteins (Sm-B, Sm-D1, Sm-D2, SmD3, Sm-E, Sm-F and Sm-G) that form a surrounding ring core domain. A further 3 U1-specific proteins, U1-70K, U1-A and U1-C bind via stem loop 1 and 2 of U1 snRNA. U1-C binding to the U1 snRNA is facilitated via the N terminus of U1-70K and the C terminus of Sm-D3 (Pomeranz Krummel et al. 2009).

The SnRNPs, U1, U2, U4 and U5 all traverse the cytoplasm for processing prior to re-entry into the nucleus (Will et al. 2001). The minor spliceosomal components U11, U12 and U4atac are also thought to be processed via the cytoplasm. Following transcription and capping within the nucleus the snRNA precursors bind an export complex to traverse into the cytoplasm. The export complex dissociates allowing the Sm proteins to bind and form the snRNP core, mediated by the survival of motor neurone (SMN) complex (Massenet et al. 2002). Following cap hypermethylation and 3' end trimming the snRNP core is transported back into the nucleus via an import complex. Once in the nucleus other snRNP associated proteins then attach e.g. U1-70K, U1-A and U1-C in the case of U1 snRNP. Final modification, processing and association between snRNPs e.g. formation of the tri-

snRNP U4/U6.U5 occurs within the nucleus in the Cajal bodies (Novotny et al. 2011). U6 is the exception in that it is transcribed by RNA Pol III and acquires its LSm (Sm-like proteins) ring whilst remaining in the nucleus (Mayes et al. 1999; Listerman et al. 2007). U6atac generation and processing occurs along similar lines.

Kinetic analysis of spliceosome assembly on a single pre-mRNA within yeast (*Saccharomyces cerevisiae*) whole cell extract has been achieved via co-localisation single-molecule spectroscopy (CoSMoS) (Hoskins et al. 2011). Spliceosomal components (U1, U2, U4/U6.U5 and the Prp19-complex) were protein tagged at the C terminus to allow different fluorophores to be attached, which could then be detected simultaneously by CoSMoS. These studies confirmed both the order and kinetics of complex binding and subsequent disassembly in the order U1, U2, U4/U6.U5 and finally Prp19. Commitment to splicing appears to increase as assembly proceeds but there was evidence that complexes may bind and dissociated multiple times prior to a commitment to splicing (Hoskins et al. 2011).

1.2.4.3 The splicing reaction

The pre-mRNA is generated following transcription and is converted to mRNA through the process of splicing. A recent study measured the half-life of pre-mRNA from over 2000 genes in mouse dendritic cells in response to lipopolysaccharide and predicted the half-life (for 95% of genes) to be between 1 to 30 minutes, with a peak of 5 minutes. The wide distribution related to a number of factors including, the presence of co-transcriptional splicing (Chapter 1.2.8.4), constantly expressed

vs. intermittently expressed genes, the number of exons and length of the transcript (Rabani et al. 2011).

Splicing of pre-mRNA occurs within the nucleus of the cell. That said, splicing has been shown to occur outside the nucleus in some situations e.g. splicing of the cytokine interleukin-1 (IL-1) occurs in the cytoplasm of anucleate platelets (Denis et al. 2005). The splicing reaction is a culmination of an intricate series of interactions principally between the pre-mRNA, the spliceosome and a number of snRNP associated and non-associated proteins (figure 1.2.4.B). With the formation of each complex numerous proteins associate and disassociate. Conformational changes may be mediated by DExD/H ATPases/RNA helicases and peptidyl-prolyl isomerases (PPIases) (Staley et al. 1998). The function of the involved proteins may be switched by post-translational modification including phosphorylation, ubiquitination and acetylation to allow further regulation (Mathew et al. 2008; Bellare et al. 2008; Kuhn et al. 2009).

The splicing choices at any given point are based on a snapshot configuration of a multitude of competing factors including splice site strength, exon and intron length and the distribution of splicing enhancers and inhibitors. Splicing choices are tissue specific and are influenced by the expression of tissue specific or ubiquitous splicing regulators (Chen et al. 2009). These factors interplay to determine splicing choices.

The process is most well characterised in yeast. At the core of the reaction is the removal of the intron and ligation of exons via two transesterification reactions at the 5′ and 3′ splice site (Staley et al. 1998). The first transesterification reaction consists of the 2-hydroxyl group of an A residue in the branch sequence attacking the phosphate forming the 3’-5’ phosphodiester bond at the 5’ splice site. The second step involves the 3-hydroxyl group of the upstream exon attacking the 3’

end of the lariat intron. Both splicing steps have been shown to be reversible in vitro (Tseng et al. 2008).

The E (early or commitment) complex forms with binding of U1 snRNP to the 5′ splice site independent of ATP. SF1 binds the branch point sequence. U2 auxiliary factors (U2AF), consisting of 65 and 35 kDa subunits, binds both the polypyrimidine tract and the AG dinucleotide at the 3′ splice site respectively. Studies indicate that U1C, a U1 snRNP associated protein, is involved in recognition of the 5' splice site via RNA-protein interactions prior to base-pairing between U1 snRNP and the 5' splice site (Du et al. 2002).

SF1, via its RNA recognition motif (RRM) domain, co-operatively binds to the U2AF65 subunit. X-ray structural examination reveals that dimerisation between the U2AF subunits occurs via an atypical RRM of U2AF35 which behaves as a peptide binding motif. This motif facilitates a protein-protein interaction with a polyproline helix of U2AF65 via reciprocal tryptophan side-chain residues (Kielkopf et al. 2001).

The U2AF35-AG interaction is stabilised by the phosphorylated RNA-associated protein, DEK which also limits binding of the U2AF65 to a polypyrimidine tract when not followed by the AG dinucleotide (Soares et al. 2006).

U2AF65 plays a role in determining splicing efficiency based on the strength of the polypyrimidine tract along with its role as a scaffold splicing factor. U2AF65 contains a pair of RRM domains, RRM1-RRM2. A strong polypyrimidine tract results in the RRM1-RRM2 taking up an open conformation leading to enhanced U2 snRNP recruitment (Mackereth et al. 2011).

The choice of splice sites and potential alternative sites has traditional been thought to be determined with the initial binding of the spliceosome, thus the E

complex is also known as the 'commitment' complex. However alternative splicing decisions may occur later than this. An upstream AG dinucleotide is chosen as a 3' splice site in preference to a downstream AG in intron 2 of the Drosphila Sex-lethal (SXL) pre-mRNA in the presence of the splicing factor SPF45 only just prior to the second step of splicing (Lallena et al. 2002).

The A complex is formed when SF1 is released from the BPS and U2 snRNP, which is thought to loosely bind to the BPS in the E complex independently of ATP, becomes stably bound sequence via ATP hydrolysis (Das et al. 2000). This process is facilitated by the DExD/H-type RNA-dependent ATPases/helicases Sub2/UAP56 and Prp5. Base-pairing between U2 snRNP and the BPS is stabilised by the 14 KDa subunit of SF3b (SF3b p14) binding to the adenine of the BPS and SF3b 155 subunit to the RRM of U2AF65 (Will et al. 2001; Gozani et al. 1998).

The complex B forms as the tri-snRNP U4/U6.U5 binds the 5′ splice site facilitated by the U5 associated splicing factor, Prp8 (human U5-220 KDa) (Maroney et al. 2000). U5 binds both the 5' and 3' exons. The 3' end of U6 and the 5' end of U2 base-pair (Valadkhan et al. 2000). U4 and U6 are bound such that the catalytic core of U6 is inactive.

The B* complex (activated spliceosome) forms with release of U1 from the 5′ splice site. Characterisation of the B* complex identified over 100 associated proteins including cyclophilins, the RNA helicases Prp22 and p68 and a number of proteins belonging to the Prp19 (NineTeen Complex or NTC) complex (Makarova et al. 2004). For the 1st step to proceed the 5' splice site, bound to the U6-GAAACA sequence is repositioned to the first step catalytic center of U6 (sequence-

ACAGAG) under the influence of the Prp28 protein and the Prp19-associated complex (Chan et al. 2003). The helicase Brr2, component of the U5 snRNP, unwinds the U4/U6 complex under the control of the associated proteins Prp8 and the GTPase Snu114 (Small et al. 2006). The U4/U6 associated proteins and those holding U5 in place disassociate. Further U6:U2 duplexes are formed.

The Prp19 complex has been shown to play a crucial role in the U4/U6.U5 snRNP interactions with the pre-mRNA. The Prp19 complex becomes associated with the spliceosome after U1 and U4 have dissociated. It mediates interactions between the 3' terminal tract of U6 and the intron near the 5' splice spite which is initiated when the 3' terminal of U6 is exposed following the destabilisation of LSm proteins bound to it (Chan et al. 2003).

The C complex signifies catalytic activity of the spliceosome. In the C1 complex U4 is released via ATP hydrolysis and U6/U2 catalyses the 1st transesterification step with cleavage at the 5′ splice site and formation of the lariat. The 5' splice site interaction with the U6 snRNA (ACAGAG) is disrupted prior to the second step (Konarska et al. 2006). Complex C2 signifies the 2nd transesterification reaction with cleavage at the 3′ splice site, exon ligation and lariat release.

Comparison of purified catalytic spliceosome C complex with B complex in man has identified differences in the associated protein content (Bessonov et al. 2008). The B complex was shown to be associated with 130 proteins in comparison to the 150 proteins with the C complex. One hundred and five proteins where found to be common between both complexes. More precisely, in association with the C complex the Prp19 complex was enriched while two multi-subunit complexes SF3a and SF3b were depleted. Thirty-seven proteins were identified to be tightly bound

in the C complex following treatment with 1M salt indicating their likely role in the catalytic activity of the C complex. These included the Prp19 complex, U5-220 (yeast equivalent Prp8), U5-116 (Snu114), U5-40K (Prp8-binding protein). Other second step factors found in the C complex and not associated with the B* complex include hPrp17, the RNA helicase DDX35, and HSP73 (Makarov et al. 2002).

The components of the spliceosome making up the catalytic core are still debated. Protein free RNA complex associated with fragments of U2 and U6 snRNA in vitro demonstrate the 2'OH group of a adenine at the branch site attacking a phosphodiester bond akin to the first step of splicing, suggesting that these snRNA make up the essential catalytic core of this RNA enzyme (Valadkhan et al. 2001). However several protein components remain near the catalytic active site, namely the U2 specific SF3b 14a/p14 and U5-specific Prp8 and may also play a role (Will et al. 2001; Abelson 2008).

Once splicing has occurred and the mRNA has been generated, the ATP-dependent DExH box RNA helicase, Prp22 mediates mRNA release (Company et al. 1991). This leaves the lariat with various remnant RNPs from the spliceosome designated the Intron large (IL) complex (Yoshimoto et al. 2009). Spliceosome disassembly is mediated by the DExH box RNA helicases, Prp43, Brr2 and the GTPase Snu114 bound to GTP (Arenas et al. 1997; Small et al. 2006). Prp43 is found in the trimeric NTR complex in association with two splicing factors Ntr1 and Ntr2. Ntr1 contains a G-patch domain (RNA binding domain) that binds Prp43 and induces its helicase activity in spliceosome disassembly (Tsai et al. 2005; Tanaka et al. 2007). Once the snRNPs are released the Intron Small (IS) complex is generated which can undergo degradation more readily than the IL complex. The lariat is linearised by the debranching enzyme DBR1 and then degraded (Moore 2002).

Figure 1.2.4.B

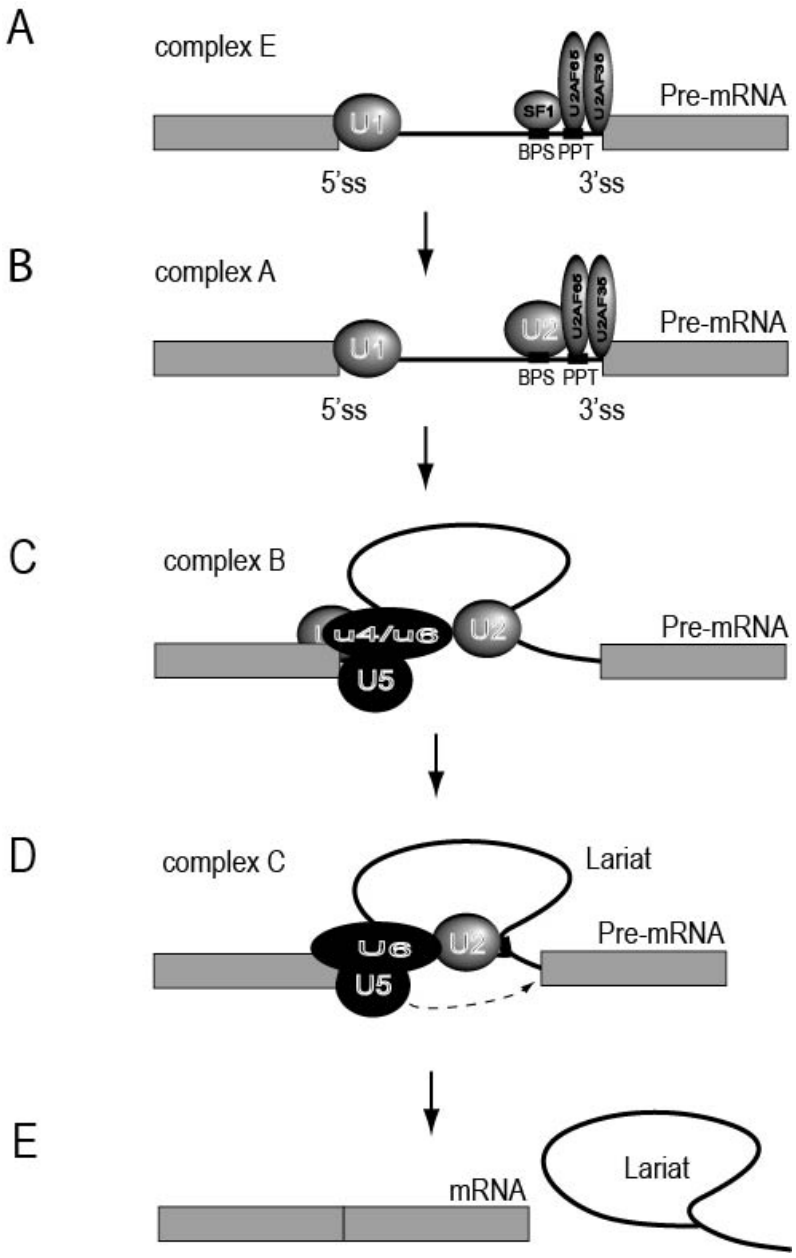


Figure 1.2.4.B The splicing mechanism. Schematic figure demonstrating the splicing mechanism. Exons are given as grey boxes and introns as black lines. Various intermediate complexes in the splicing reaction are shown. These include complexes E, A, B and C. Other intermediate complexes and splicing factors are involved in the splicing process but have been omitted for clarity.

1.2.5 Splicing enhancers and repressors

1.2.5.1 Cis- acting elements.

Sequences within the pre-mRNA are known to influence the choice between splice-sites to allow accurate and reproducible splicing at the latter. Exonic and intronic splicing enhancer sequences (ESE and ISE, respectively), which tend to be purine rich, have been characterised, and in general act by binding SR proteins which interact with the spliceosome to enhance the use of particular splice sites (Liu et al. 1998; Fairbrother et al. 2002; Zhang et al. 2004). Intronic and exonic splicing silencer sequences (ISS and ESS, respectively) have also been identified on functional and bioinformatic grounds (Wang et al. 2004; Zhang et al. 2004; Yu et al. 2008). Using an in vitro SELEX approach with in vivo confirmation, ISS and ESS sites were identified on the basis of selecting an upstream weak 5' splice site for U1 snRNP binding in preference to a downstream consensus 5' splice site (Yu et al. 2008). The silencer sites often interact with members of the family of heterogeneous nuclear ribonucleoproteins (hnRNPs) including hnRNP A1, polypyrimidine tract binding protein and hnRNP H, to suppress the use of particular splice sites (Chen et al. 1999; Del Gatto-Konczak et al. 1999).

1.2.5.2 Trans-acting elements.

Trans-acting factors include splicing enhancers in the form of SR proteins (see section 1.2.5.3.) and splicing silencers such as the hnRNP family. These sequences interact with cis-acting elements as described above and enhance or inhibit the splicing process via several different mechanisms. The balance between enhancing factors e.g. SF2/ASF and silencing factors e.g. hnRNP A1 determines the overall outcome of splicing (Mayeda et al. 1992; Bai et al. 1999). The effects of

trans-acting elements on splicing are more complex, however, than simply those that enhance and those that inhibit splicing. For instance, the NOVA1 protein acts as a splicing enhancer by causing exon 9 inclusion when bound to an ISE in the GABA A receptor γ2 and as a splicing repressor for exon 4 of its own pre-mRNA when bound to an ESS (Dredge et al. 2005).

1.2.5.3 Splicing enhancers.

The SR proteins are the major family of trans-acting splicing factors which play a role both in constitutive and alternative splicing. Further studies have also identified that SR proteins have other diverse roles in mRNA nuclear export, translation and nonsense mediated decay (Caceres et al. 1998; Sanford et al. 2004; Zhang et al. 2004). These proteins consist of arginine/serine rich repeat domains also known as RS domains (RS or SR denotes single letter notation for arginine - 'R' and serine - 'S') at the C-terminus and a number of RRMs at the N-terminus. The RRM domain, which can bind to specific RNA sequences, specifically interacts with the pre-mRNA while the RS domain can mediate both protein-protein and protein-RNA interactions. Splicing enhancers may mediate their effects by recruiting relevant components of the spliceosome, other splicing factors or by antagonising silencer elements (Zuo et al. 1996; Blencowe 2000; Graveley et al. 2001; Kan et al. 1999). The RS domain also plays a role in SR protein cellular localisation via interactions with transportin-SR, the SR protein nuclear import receptor (Lai et al. 2000).

Post-translational modification of SR proteins can modify their effect. The SR protein SRp38, a splicing repressor in the dephosphorylated state becomes an activator in the phosphorylated state by recruiting and stabilising U1 and U2 snRNP (Shin et al. 2002; Feng et al. 2009).

SR proteins appear to be regulated in part by ultraconserved genomic elements,

defined as regions of at least 200 nucleotides of perfect identity between the human, mouse and rat genomes, found within their coding genes (Bejerano et al. 2004). These elements code either alternative poison cassette exons or the 3' UTR of the gene and allow transcript degradation, postulated to be a means of self regulation (Lareau et al. 2007).

SR-related proteins, a further group of splicing factors which contain an RS domain but may not have an RRM domain also play roles in splicing and other cellular functions (Boucher et al. 2001).

1.2.5.4 Splicing silencers

The best characterised family of silencing factors is the Heterogeneous nuclear ribonucleoproteins (hnRNP) family of proteins which work by interacting with negative cis-acting regulators. HnRNP were originally defined as proteins associated with RNA in the cell nucleus. These are a heterogeneous family of proteins with marked structural and functional diversity. The mammalian hnRNP family consists of over 20 distinct polypeptides (Swanson et al. 1988). Like SR proteins, hnRNP proteins consist of one or more RNA-binding domains associated with an auxiliary domain which is involved in protein-protein interactions. Mayeda and colleagues established that hnRNP A1 is able to modulate 5′ splice site selection directly (Mayeda et al. 1993). HnRNP A1 and hnRNP H (PTB) are the best characterised members of the family which play a clear role in splicing repression. Despite this, there is evidence that they positively modulate the splicing process depending on their binding position (Xue et al. 2009). Analysis of PTB-RNA interactions across the genome using high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP or CLIP-seq) demonstrated that PTB binding near constitutive splice sites enhances the inclusion

of adjacent alternative exons. HnRNP proteins also appear to play roles in transcription, telomere biogenesis, polyadenylation, translation, RNA editing and mRNA stability (Krecic et al. 1999).

Putative mechanisms have been invoked to explain the action of splicing silencers (figure 1.2.5.A) (Cartegni et al. 2002).

1. Occlusion - Splicing silencers such as hnRNP A1, hnRNP and hnRNP H may directly compete with enhancer factors to occlude binding sites in close apposition to each other in a mutually exclusive manner resulting in an inhibition of splicing.
2. Nucleation and co-operative binding - ESS sites may exert an inhibitory effect on distant ESE sites by binding an inhibitory factor e.g. hnRNP which nucleates the co-operative polymerisation of further inhibitory factors to displace or inhibit enhancer factors binding (Zhu et al. 2001).
3. Loop-out - Silencer splicing factors, bound to ISS elements either side of a given exon may dimerise and loop-out the exon resulting in an inhibition of splicing (Blanchette et al. 1999). The binding of pre-mRNA to the RNA binding domains of PTB will tend to result in ‘looping-out’ of a single strand of RNA (Oberstrass et al. 2005).
4. Repression of pseudo and cryptic splice sites - hnRNP binding sites have been identified in close proximity to pseudo and cryptic splice sites and are thought to tonically inhibit their use thereby allowing constitutive exons to be spliced.

Figure 1.2.5.A

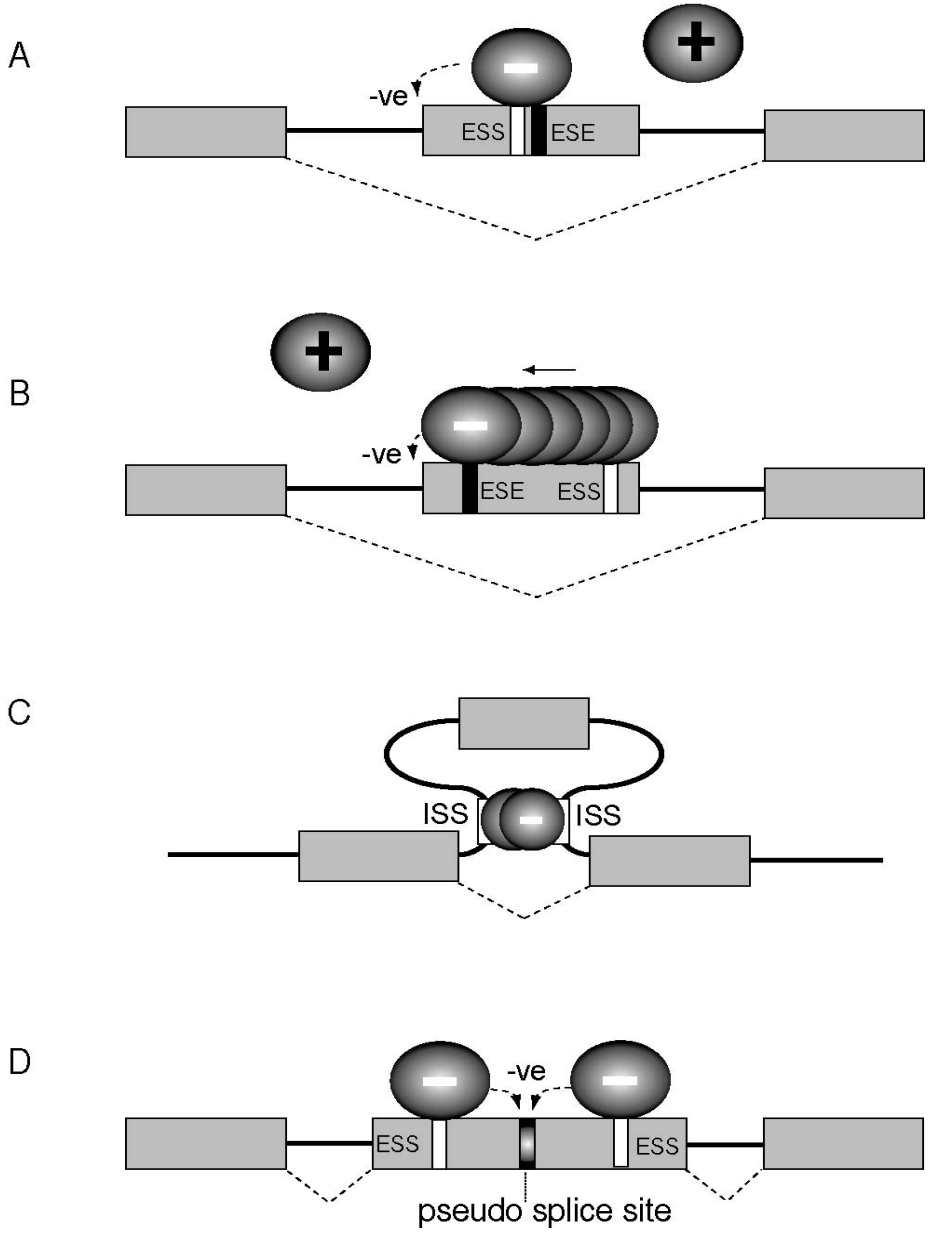


Figure 1.2.5.A Mechanisms of splice silencing. Schematic diagram shows a portion of pre-mRNA. Exons are shown as grey boxes and introns are denoted as black lines. Cis-acting enhancer sequences (ESE and ISE) and inhibitory sequences (ESS and ISS) are labelled accordingly. Splicing factors are given as shaded circles with a **+** denoting a splicing enhancer and **–** denoting an inhibitory factor. The four figures illustrate mechanisms of splice silencing **A.** Silencing by occlusion. **B.** Silencing by nucleation and co-operative binding. **C.** Silencing by exon looping. **D.** Silencing pseudo splice sites by associated ESSs. Modified from Cartegni et al. 2002 {Cartegni, 2002 #709}.

1.2.6 Intron and exon definition.

Two potential mechanisms exist to explain how the spliceosome identifies intron-exon boundaries in the pre-mRNA. Exon definition postulates that pre-spliceosomal components identify and bind to pairs of splice sites around an exon (5' splice site and the upstream 3' splice site). Intron definition suggests that pairs of splice sites around introns are defined initially. Whilst clearly both processes occur, the dominant mechanism will, in part, be determined by the ability of the cross intron or cross exon complex to span the divide. The predominant mechanism in higher order eukaryotes, in which introns tend to be significantly longer, appears to be exon definition (Robberson et al. 1990). 5′ splice site mutational studies confirm this by demonstrating more proximal exon skipping rather than distal intron retention in keeping with exon definition models (Nakai et al. 1994).

Inherent exonic features which determine splicing choices include exon length and splice site strength. The inherent strength of splice sites is an important factor in splice site choice (figure 1.2.7.A) which may be estimated using a consensus scoring matrix (Shapiro et al. 1987).

The majority of exons are between 50-300 nucleotides in length with an average of 120 nucleotides (Dominski et al. 1992; Ast 2004). Evidence for the proposition that there is an optimum length for an exon includes the fact that only 1% of exons are longer than 400 nucleotides (Berget 1995); exons shortened to less than 50 nucleotides are more likely to be skipped compared to longer exons (Dominski et al. 1991); constitutive exons expanded to greater than 300 nucleotides are increasingly skipped (Robberson et al. 1990) and cryptic splice sites often emerge very near to mutated splice sites, again suggesting that there is an optimum range for exon length. Very short constitutively spliced exons do exist e.g. NCAM has a 3 nucleotide exon (Santoni et al. 1989).

Figure 1.2.7.A

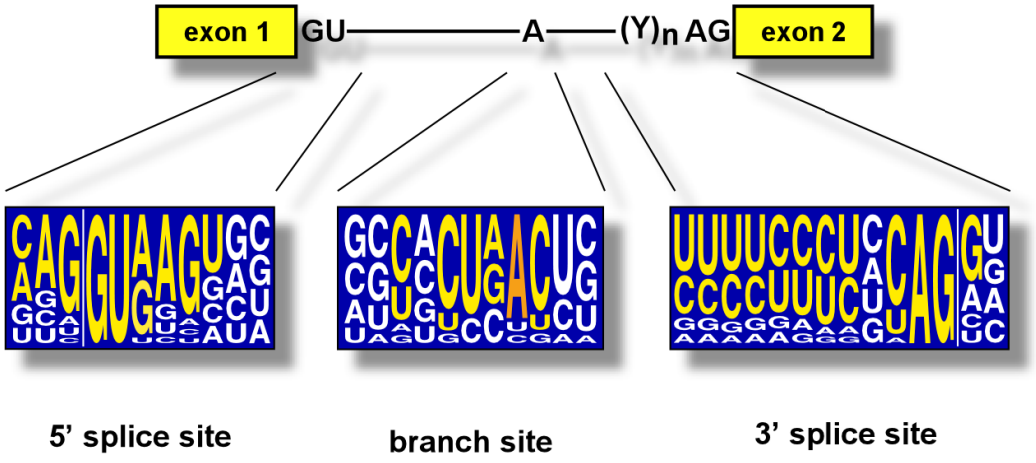


Figure 1.2.7.A Splice site strength. Schematic diagram of pre-mRNA with consensus sequence splice sites. Exons are represented as boxes and the intron as a black line. The three boxes below represent the frequency with which nucleotides appear in a given position at the 5′ splice site, branch point sequence and 3′ splice site. Adapted from Cartegni et al. 2002 {Cartegni, 2002 #709}.

Less favourable exon lengths may be overcome by optimising splice site strength and near-by enhancer sequences (Black 1991; Dominski et al. 1992).

Intron definition is considered more likely to occur when the length of the intron is less than 200 nucleotides (Fox-Walsh et al. 2005). Highly expressed genes tend to have shorter introns than genes that are expressed less (Castillo-Davis et al. 2002). Constitutively splice exons which are expanded until they are skipped may again be constitutively spliced if the surrounding introns are shortened (splicing by intron definition) (Chen et al. 1994).

Exon definition by spliceosomal components must convert to intron definition prior to intron excision. Though the mechanisms for this are poorly understood, insights into this process come from Schneider et al. using a trans-splicing approach to demonstrate that cross exon complexes include not only the U1 and U2 snRNP but also the tri-snRNP U4/U6-U5, which is able to positionally convert from a cross exon defined complex to bind the upstream 5' splice site across the intron (Schneider et al. 2011). Binding to the 5' splice site occurs via a U6 interaction suggesting a mechanism to convert cross exon to cross intron complexes. U6 can bind directly to the 5' splice site in the absence of U1 snRNP to form a B complex suggesting it is possible to bypass the need for a prior A complex (Schneider et al. 2011).

The conversion from exon to intron definition can also be used as a point of regulation. The RNA-binding motif protein 5 (RBM5) appears to regulate splicing of exon 6 of the apoptotic factor FAS by inhibiting the conversion of the exon defining to intron defining complex (Bonnal et al. 2008). This is achieved by inhibiting the binding of the tri-snRNP U4/U6-U5 to the flanking introns leading to exon 6 exclusion.

Splicing around exon 4 of the CD45 pre-mRNA is preceded by the formation of an A-like exon definition complex (AEC) across exon 4. The binding of hnRNP L to an exonic splicing silencer (ESS) in exon 4 appears to inhibit cross intron definition by U1 and U2 snRNP to inhibit splicing (House et al. 2006).

Several mechanisms have been put forward to explain how the spliceosome might link splice sites across huge introns. The Ultrabithorax gene of Drosophila contains a 74 KB intron and provides an example of 'recursive' splicing whereby consecutive sections of intron are spliced out between the upstream 5' splice site and

sequential downstream 3' splice site (''ratcheting point'' or RP-sites) within the intron. Each time a section of intron is spliced out the spliceosome forms repeatedly using the upstream 5' splice site and the next 3' splice site down the intron until it splices an entire long intron (Hatton et al. 1998).

Recursive splicing is not thought to be a significant mechanism to manage long introns in mammals however. Bioinformatic studies in vertebrates suggest an alternative strategy using short and long interspersed element repeats (SINE and LINE) which are able to form stems to loop out large lengths of intronic sequence thereby bringing the 5' and 3' splice site closer together to enhance intron definition and splicing (Shepard et al. 2009).

1.2.7 RNA secondary and tertiary structure

RNA secondary and tertiary structure can influence the rate and extent of alternative splicing. The secondary structure of RNA refers to sequences of classical Watson-Crick base-paired and unpaired ribonucleotides and may be represented in a two dimensional fashion. It includes structures such as the hairpin loop, bulges, internal loops and stem loops (Chastain et al. 1991). RNA tertiary structure is the three dimensional interaction of secondary structure elements and includes the ribose zipper, coaxial stacking, kissing hairpin loops and tertraloops (Hermann et al. 1999).

In vivo studies have demonstrated that secondary structure formation following transcription can affect alternative splicing choices (Eperon et al. 1988). Secondary structure may either enhance exon skipping or inclusion by bringing relevant splice sites into close apposition or by obscuring sites from spliceosome recognition. Analysis suggests that splicing enhancer and inhibitory motifs tend to appear in predicted single strand rather than double stranded RNA sequences where they

are able to exert an effect (Hiller et al. 2007). Indeed the effects of enhancer and inhibitory motifs are modified according to the secondary structure conformation they are found within (Buratti et al. 2004).

The effects of secondary structure on alternative splicing were demonstrated by Graveley in the 48 potential isoforms of exon 6 of the DSCAM gene in Drosophila melanogaster (Graveley 2005). Using comparative genomics to look at genomes of 16 different insects, two types of conserved sequence were identified. The docking sequence appeared upstream of all exon 6 variants and selector sequences appeared upstream of each potential exon 6 variant. The presence of complimentary sequences in the docking sequence with each of the selector sequences provides a mechanism where-by any of the exon 6 variants may be selected in a mutually exclusive secondary structure by binding of the docking sequence to a given selector sequence.

1.2.8 Environmental cues to splicing

As part of the cellular response to environmental cues, modulation of splicing by intracellular and extracellular changes in the environment might be expected. Evidence is now accumulating for the influence of environmental cues on the process of splicing.

Changes in oxygen tension have been shown to regulate cellular alternative splicing in the tumour antigen gene MGEA6 (Webby et al. 2009). The splicing factor U2AF65 undergoes lysyl-hydroxylation by a homologue of the oxygen sensitive HIF asparaginyl hydroxylase, the dioxygenase Jumonji domain-6 protein (jmjd6) (Webby et al. 2009). Jmjd6 knockdown resulted in altered splicing of the MGEA6 gene (exon 19 skipping) linking environmentally triggered cues to changes in splicing.

Another example of an environmental trigger affecting splicing comes from the splicing repressor SRp38 and cellular heat shock. Heat shock results in dephosphorylation of SRp38 which interferes with the U1 snRNP-5' splice site interaction and causes a reduction in splicing efficiency (Shin et al. 2002).

Extracellular cues transduced via extracellular-signal-regulated kinase/ mitogen-activated protein kinase (ERK/MAP-kinase) within the RAS pathway may influence the alternative splicing of exon v5 of the adhesion transmembrane molecule, CD44 (Matter et al. 2002). Sam68, an RNA-binding protein binds to exonic splicing regulatory elements resulting in exon v5 inclusion when phosphorylated by ERK.

Riboswitches are cis-acting RNA elements found predominately in bacteria but also plants and fungi which respond to a metabolite and can regulate not only transcription and translation but also alternative splicing. Thiamine pyrophosphate

(TPP) - responsive riboswitches in the filamentous fungus Neurospora crassa have been shown to both activate and repress gene expression by modulating alternative splicing (Cheah et al. 2007). In the case of the N-myristoyltransferase 1 (NMT1) pre-mRNA which itself is involved in thiamine metabolism, the riboswitch, bound to thiamine, unmasks an alternative downstream 5' splice site. This results in the generation of the alternative isoform, I-2 which competes with the constitutive isoform for translation.

1.2.9 Chromatin structure and histone modification

The nucleosome consists of a number of histone proteins bound within a DNA strand and constitutes the basic building block of chromatin. The binding of the nucleosome to exons and introns influences both transcription and splicing. Several lines of evidence support the notion that chromatin structure plays a role in alternative splicing. DNA topoisomerase I (Top1), involved in relaxing DNA supercoiling prior to transcription by RNA polymerase II, is known to phosphorylate SR proteins including ASF/2SF, essential in mediating exon skipping decisions (Rossi et al. 1996). Furthermore the SR protein B52 has been shown to play an essential role in localising the Drosphila DNA topoisomerase I to active transcription sites as well as in mRNA release following transcription (Juge et al. 2010). Genome wide mapping of chromatin structure identifies that nucleosomes are placed non-randomly (nucleosome phasing) in the vicinity of exons rather than introns, constitutive rather than alternatively spliced exons and are less frequent at the site of pseudoexons (Spies et al. 2009; Tilgner et al. 2009; Kolasinska-Zwierz et al. 2009).

There are many examples to illustrate the association between chromatin structure and alternative splicing outcomes. Inclusion of the exon E33/EDI of the fibronectin

pre-mRNA is dependant on chromatin arrangement (chromatinisation) of the plasmid (Kadener et al. 2001). Methylation of histones drives interactions with splicing factors via chromatin adaptor proteins to alter splicing choices. Splicing factors have been shown to bind to the nucleosome via histone modifications. Histone modification, in the form of trimethylation of lysine 36 of histone H3 (H3K36me3) leads to PTB binding via the chromatin adaptor proteins MRG15 and a reduction of the inclusion of a PTB-dependent exon (exon IIIb) in the human fibroblast growth factor receptor 2 (FGFR2) gene (Luco et al. 2010).

1.2.10 Natural antisense transcripts (NAT)

Recent genome wide studies of the eukaryotic transcriptome have focused attention on the diverse non-coding RNAs (ncRNA) which include natural antisense transcripts (NAT) generated from the antisense (second) DNA strand. NATs appear to play a role in regulating gene expression at all levels of transcriptional and post-transcriptional control including alternative splicing. NATs may bind to sense pre-mRNA transcripts and block spliceosome and splicing factor interactions.

The expression of the thyroid hormone receptor gene, ErbAa in B lymphocytes has been shown to be regulated by antisense RNA strands (Hastings et al. 1997).

NATs have been shown to down-regulate E-cadherin expression by binding the mRNA of Zeb2, a transcriptional repressor of E-cadherin. NAT binding leads to intron retention at the 5'UTR which contains a key ribosomal site required for Zeb2 protein expression (Beltran et al. 2008).

1.2.11 Co-transcriptional splicing

Co-transcriptional splicing is the process whereby splicing occurs concurrently with transcription and is mediated by RNA Pol II linking the splicing machinery with transcriptional elements (Perales et al. 2009). Post-transcriptional splicing does occur, predominantly with introns towards the 3' end of the transcript (Bauren et al. 1994). Evidence for co-transcriptional splicing comes from several sources; the observation that chorion gene transcripts in Drosophila shortened (i.e. spliced) whilst still attached to chromatin and RNA Polymerase (Osheim et al. 1985); the presence of introns in an RNA transcript increased transcriptional efficiency (Brinster et al. 1988) and the rate of RNA Pol II transcription significantly alters the rate of alternative splicing (de la Mata et al. 2003). Chromatin immunoprecipitation (ChIP) is a technique in which a protein of interest is cross-linked, using agents such as formaldehyde, to target DNA which may subsequently be retrieved and sequenced to establish interacting partners. ChIP studies have confirmed that elements of the spliceosome (U2, U5 snRNP, Prp19) assemble on intron containing genes prior to the completion of its processing by RNA Pol II (Gornemann et al. 2005).

Aubeouf et al. demonstrated that the addition of transcriptional co-regulators e.g. the DEAD-box RNA helicase p72 and the heterogeneous nuclear ribonucleoprotein-like protein CoAA (coactivator activator), can alter splicing in a transfected minigene construct (CD44) in which a steroid responsive promoter is inserted (ERE-CD44) in response to cellular stimulation with oestrogen (Auboeuf et al. 2002).

Two models of co-transcriptional splicing exist, the recruitment model and the kinetic model. The recruitment model suggests that alternative splicing is either

enhanced or repressed by the interaction between RNA Pol II and splicing factors. De la Mata and Kornblihtt demonstrate that the C terminal domain (CTD) of RNA Pol II interacts with SRp20 to inhibit the inclusion of the alternative EDI exon of the fibronectin gene (de la Mata et al. 2006).

The kinetic model postulates that the likelihood of an alternative exon being included depends of the rate of transcription as the result of RNA Pol II. Evidence for this postulation comes from in vivo experiments in which a splice site within a stem structure is utilised depending on size of an intervening loop suggesting that the rate of transcription determines the secondary structure formation possible and ultimately the use of a given splice site (Eperon et al. 1988). Below a threshold loop size secondary structure is able to form and sequester the splice site. Above the threshold loop size the stem formation is delayed allowing the splice site to be utilised. The splice site was poorly utilised in vitro irrespective of the loop size as the transcript is pre-formed and there is no impediment to secondary structure formation. Another study demonstrated that the degree of inclusion of alternative exons with weak splice splices increased when transcription speeds were decreased as the result of a point mutation in RNA Pol II (de la Mata et al. 2003).

There are multiple examples of interactions between the process of transcription and splicing. TAT-SF1, an elongation factor, interacts with U2 snRNP to stimulate splicing and concurrently promote transcriptional elongation via interactions with a component of the transcription elongation complex, P-TEFb (Fong et al. 2001). Interactions between U1 snRNP and several splicing regulatory proteins allows RNA Pol II to transcribe more efficiently than with other polymerases e.g. T7 phage which do not (Das et al. 2007). The splicing factor SC35 is known to promote more

efficient transcription elongation (Lin et al. 2008). Long introns may be degraded by exonucleases as transcription takes place or so called co-transcriptional cleavage of introns. Splicing remains unaffected as the upstream exons is held by the polymerase elongation complex (Dye et al. 2006).

1.2.12 Polyadenylation

Polyadenylation is the addition of a multiple adenine nucleotides at the 3' end of the transcript after transcription and marks the generation of a mature mRNA. The process of polyadenylation and splicing are again linked. Polyadenylation signals enhance with terminal intron definition and binding of the 3' splice site processing machinery. Both U2AF and components of the U2 snRNP binding the terminal 3' splice site interact directly with the poly(A) complex component, the 73-kDa subunit of cleavage and polyadenylation specificity factor (CPSF-73) (Kyburz et al. 2006). The U1 snRNA binding to the 5' splice site has been shown to suppress potential upstream intronic polyadenylation signals. This was demonstrated using whole viral genome studies with antisense knockdown of U1 snRNA (Kaida et al. 2010).

Cleavage at the poly(A) site at the end of an mRNA transcript is mediated by a component of the poly(A) complex with endonuclease activity, CPSF-73 (Mandel et al. 2006). Degradation of RNA transcripts beyond the poly(A) tail comes via a complex with 5'-3' exonuclease activity including polypeptides Rtt103, Rat1 and Rai1 in yeast (Kim et al. 2004). Higher eukaryotes require additional mechanisms to cleave and degrade the RNA transcript beyond the poly(A) tail. These include transcriptional pauses allowing the 5'-3' exonuclease Xrn2 (Rat1) to degrade the transcript and the presence of a termination element approximately 1Kb downstream of the poly(A) site which allows co-transcriptional cleavage (Kaneko et al. 2007; West et al. 2004).

1.2.13 mRNA Export

Nuclear export of the resulting mRNA from the nucleus to the cytoplasm occurs in 3 distinct stages; firstly mRNP formation; secondly, targeting and translocation through the nuclear pore complex (NPC) and finally mRNP release and disassembly (Carmody et al. 2009). The mRNA is transported in association with proteins involved in RNA processing and export as a messenger ribonucleoprotein (mRNP). Some of these are a part of the family of hnRNPs. Processing, from transcription to nuclear export are closely linked (Kohler et al. 2007). Surveillance mechanisms are in place to ensure nuclear processing is complete prior to nuclear export. Transfer is as part of a cargo-carrier module via a NPC. NPCs are conduits between the nucleus and cytoplasm by which macromolecules are transferred and are made up of nucleoporins.

Several complexes including the THO/TREX complex and TREX 2 in man play a central role in mRNA export but also appear to be linked to transcription and splicing (Strasser et al. 2002; Luna et al. 2009). The THO/TREX complex consists of the heterotetrameric TREX ('transcription/export') complex and the THO complex consisting of the mRNA export factor Yra1 (REF/Aly in mammals) and the DEAD box protein Sub2 (UAP56 in mammals). Sub2 also plays a role in spliceosome assembly (Shen et al. 2008). In man TREX is bound to the 5' end of the mRNA via a combination of cap-binding proteins, Yra1 (Aly/REF) and exon junction complex (EJC) factors (Cheng et al. 2006).

NXF1-NXT1/TAP-p15 (Mex67-Mtr2 in yeast), a conserved export carrier and TREX 2 bind the mRNA to the NPC (Gruter et al. 1998; Jani et al. 2012). This may occur directly or via several proteins including Yra1 (Aly/REF), Sub2 (UAP56) and the RNA-binding protein Np13 in S. Cerevisiae (Lund et al. 2005). SR proteins e.g.

SRp20, 9G8 and ASF/SF2 may also serve as export adaptors to export receptors such as NXF1-NXT1/TAP (Huang et al. 2005).

The DEAD box helicase DpB5 initiates mRNP disassembly on the cytoplasmic side of the NPC following activation by other factors involved in translocation of mRNA including the NPC-associated Gle1 and inositol hexakisphosphate (InsP6) (Alcazar-Roman et al. 2006). These factors also appear to play significant roles in protein translation and termination.

1.2.14 Translation

Recently spliced mRNAs are able to engage the translational machinery more efficiently leading to a higher yield of protein than equivalent cDNA of the same transcript suggesting a facilitative link between splicing and the translational machinery (Le Hir et al. 2003). This appears to be mediated via the EJC bound ribosomal complex, SKAR (S6 kinase 1 Aly/REF-like target), a target for the 40S ribosomal protein S6 kinase 1 (S6K1) which enhances translation initiation at the 5' cap of the mRNA (Ma et al., 2008). S6K1 is also recruited by other splicing factors such as ASF/2AF to encourage translation.

Wang et al demonstrated that translational nonsense codons are able to influence alternative splicing (Nonsense-associated altered splicing). Splicing rates of a T cell receptor β (TCR β) 'alternative mRNA', which in fact skips the mutation that generates the nonsense codon, are increased in its presence (Wang et al. 2002).

1.2.15 RNA surveillance and degradation

A pre-mRNA may be mis-spliced as the result of a mutation in a splice site or a splicing regulatory element. A number of surveillance and degradative mechanisms are in place, in both the nucleus and cytoplasm to identify and degrade mis-spliced transcripts. Nonsense mediated decay (NMD) remains the noted degradative pathway but other NMD independent pathways exist.

1.2.15.1 Nonsense-mediated decay

Nonsense-mediated decay is a RNA surveillance mechanism via which RNA is targeted for degradation if it contains a premature termination codon (PTC) significantly upstream from an exon-exon junction (Maquat 2004). NMD occurs during the pioneer round of translation in mammals and provides not only a means to degrade prematurely truncated transcripts, generated as the result of mutations, but also plays a routine role in the regulation of normal non-mutated transcripts. Again, elements of the splicing machinery are inextricably linked to the process of NMD. The splicing factor ASF/2AF appears to influence the degree of NMD (Gudikote et al. 2005). Indeed a number of splicing factors self-regulate their own expression by inducing NMD when in excess (McGlincy et al. 2008).

NMD is triggered when a PTC is found more than 50 nucleotides upstream of the last exon-exon junction bound by an EJC (Chang et al. 2007). The EJC consists of 4 proteins, namely; eIF4AIII; Barentsz [Btz]; Mago and Y14 which are associated with the mRNA 20-25 nucleotides upstream of the exon/exon junction (Bono et al. 2006). On encountering a PTC significantly upstream of the EJC on an mRNA transcript, the ribosome undergoes a translational pause. Eukaryotic translation release factors eRF1 and eRF3 bind the RNA helicase Upf1 and the PIK-related

protein kinase SMG-1 to form the SURF complex (Kashima et al. 2006). The SURF complex binds the EJC, precipitating phosphorylation of Upf1 by SMG-1. Phosphorylated Upf1 recruits the serine/threonine-protein kinase SMG-5, SMG-6, SMG-7 and leads to both translational repression by inhibiting conversion to the transcriptionally active 80S ribosome complex and recruitment of decay factors Dcp1a, Xrn1, and Rrp4 necessary for NMD (Isken et al. 2008). Degradation via the NMD pathway is mediated by the 5'-3' exonuclease Xrn-1, following decapping at the 5' end of the transcript and 3'-5' nucleases following deadenylation at the 3' end of the transcript (Lejeune et al. 2003; Chen et al. 2003). SMG-6 appears to have endonuclease activity and may also contribute to degradation (Huntzinger et al. 2008).

1.2.15.2 NMD independent surveillance/degradation

Within the spliceosome, the associated DExD/H box ATPases Prp5p, Prp16p, and Prp22p play a proofreading role (Smith et al. 2008; Koodathingal et al. 2010; Burgess et al. 1993; Xu et al. 2007). In the case of Prp16p, proof-reading of the 5' splice site cleavage is determined by the rate of 5' cleavage (sub-optimal 5' splicing occurs at a slower rate). Stalled transcripts were discarded with the assistance of the DExD/H box ATPase Prp43 (Koodathingal et al. 2010).

Proteins such as the yeast Mlp1p/Mlp2p may assist in the process of surveillance and degradation, acting as retention factors which hold unspliced transcripts awaiting nuclease degradation (Dziembowski et al. 2004; Galy et al. 2004).

Non-stop decay (NSD) and no-go decay (NGD) are two further degradative pathways. NSD is a process in which mRNA transcripts, attached to ribosomes, are released and diverted towards exosomal degradation if no stop codon is identified

(Vasudevan et al. 2002). If ribosomal translational of an mRNA is impeded due to a defect in the mRNA then the transcript is said to undergo NGD. In NGD the mRNA transcript is snipped at the defect site. Ribosome release is mediated by two proteins, Dom34p and Hbs1p and nucleases are recruited to degrade the mRNA fragments (Doma et al. 2006).

1.2.15.3 Endoribonucleases and exoribonucleases

Endoribonucleases and exoribonucleases play a major role in degradative pathways. Exoribonucleases mediate either 5'-3' decay or 3'-5' decay of mRNA which requires prior deadenylation of the poly(A) tail and 5' decapping of the transcript via a decapping protein complex (Fenger-Gron et al. 2005). 5'-3' exoribonucleases in S. cerevisiae include the predominantly cytosolic Xrn1p/Kem1p, involved in degradation of decapped mRNA and the predominantly nuclear Rat1p/Xrn2p and its activating partner Rai1/Dom3Z (Xiang et al. 2009). 3'-5' mRNA decay is instigated by the exosome, a protein complex found both in the nucleus and cytoplasm. The exosome has both exoribonuclease and endoribonuclease activity and plays a key role in RNA surveillance. Enzymatic activity of the complex is derived from the hydrolytic nuclease, Rrp44/Dis 3 which harbors both exoribonuclease and endoribonuclease activity via its PIN domain (Dziembowski et al. 2007; Schneider et al. 2009). The human auto-antigen PM/Scl100kDa, a homologue to the S. cerevisiae protein Rrp6, is associated with the exosome in the nucleus and cytoplasm and also confers 3'-5' exoribonuclease activity to the exosome (Allmang et al. 1999).

Endoribonucleases degrade RNA via cleavage of phosphodiester bonds within a transcript. Exosome associated enzymes with endoribonuclease activity including Rrp44/Dis3 and Swt1, allowing endoribonucleic followed by exoribonucleic

degradation in a co-ordinated fashion (Schneider et al. 2009; Skruzny et al. 2009). Endoribonucleases can be broadly divided into those involved in degradation pathways which are sequence specific including siRNA and miRNA mediated decay (DICER and AGO2) or those involved in pathways which are not sequence specific e.g. inositol requiring enzyme-1 (IRE1), RNase L, aldolase C, activator of RNA decay (ARD-1), Ras-GTPase activating protein SH3 domain-binding protein (G3BP) and apurinic⁄apuridinic endonuclease 1 (APE1) (reviewed in Li et al. 2009).

1.2.15.4 RNA interference

RNA interference is a cellular process whereby mRNA is degraded and silenced by complementary double stranded RNA. RNA interference includes silencing via small interfering RNA (siRNA), microRNA (miRNA) and piwi-interacting RNA (piRNA) providing a further mechanism for genetic regulation.

siRNA is a short sequence (21-25 nucleotides long) cleaved from a double stranded RNA molecule which may be generated endogenously or synthetically. Endogenous siRNAs are generated via cleavage of longer dsRNA by the nuclease enzyme DICER followed by separation of the strands by a helicase. The guide strand is incorporated into the RISC (RNA-induced silencing) complex. The guide strand then binds to its target mRNA and induces cleavage by the catalytic component of the RISC complex, Argonaute (Liu et al. 2004).

miRNAs are short non-coding RNAs (~21nucleotides) which bind to complimentary RNA sequences leading to transcript degradation or repression of translation. miRNAs are derived from genome-coded precursors known as primary miRNAs (pri-miRNAs) generated by RNA Pol II. These precursors fold to create intramolecular hairpins which are then processed by a member of the RNase III family, Drosha, to generate the stem-loop miRNA precursor (pre-miRNA). The pre-

miRNA is then transported to the cytoplasm to be processed by DICER via the common pathway with siRNA. miRNAs have been shown to play a role in the regulation of alternative splicing e.g. miR-133 suppresses the production of PTB2 while miR-1 suppresses both PTB1 and PTB2 which play an important role in repressing the splicing of muscle specific exons until myotube differentiation takes place in precursor muscle cells (Boutz et al. 2007).

1.2.16 Pseudosplice sites and pseudoexons.

Pre-mRNA contains sequence elements defining the 5′ and 3′ ends of introns and the branch point. In mammals, the consensus sequences of these elements are highly degenerate (Shapiro et al. 1987). As a result, matches to these consensus sequences are highly prevalent throughout the genome (Fairbrother et al. 2000) and outnumber genuine splice-sites by an order of magnitude in the human hypoxanthine guanine phosphoribosyl transferase (hprt) gene (Sun et al. 2000). These pseudo splice-sites or cryptic splice sites are not employed in wild type splicing. How the spliceosome is able to differentiate the few genuine splice sites from numerous pseudo splice sites is as yet unclear. Most of the pseudo splice sites in the human genome reside in introns, by virtue of their long lengths (Fairbrother et al. 2000; Sun et al. 2000).

A pseudoexon is a sequence between 50 to 200 base pairs within an intron which is bound by a potential 3′ pseudo splice site in association with a downstream 5′ pseudo slice site (Sun et al. 2000). Therefore very large exons contain many pseudoexons. How these large exons are spliced intact again remains unclear.

1.2.17 Splicing and disease

Whilst splicing provides an additional level of control for gene expression, the corollary is an increased susceptibility to disease when the process becomes misregulated. It has been estimated that aberrant splicing accounts for the defect in genetic disease caused by point mutations in up to 50% of cases (Teraoka et al. 1999; Ars et al. 2000). Endocrine disease states as a consequence of disruption of splicing are numerous. Disease causing mutations may affect splicing in a number of ways:

1. Primary splicing defects - mutations may cause primary splicing defects directly in the given gene.
2. Secondary splicing defects - mutation not directly resulting in a splicing defect in the given gene but which alter splicing in a secondary fashion. Mutations in genes coding trans-acting components of the splicing machinery may result in this type of defect.
3. Tertiary splicing defects - mutations which cause disease due to defects in splicing by mechanism other than primary or secondary splicing defects.

1.2.17.1 Primary splicing defects.

Primary splicing defects result from cis-acting mutations that alter splicing. These mutations may cause disease in one of several ways including; mutating a constitutive splice site; creating a cryptic splice site; altering secondary structure such that splicing is affected; altering enhancer or silencer elements or generating premature stop codons and thus stimulating nonsense-mediated decay (NMD) or nonsense-associated altered splicing (NAS) (Dietz et al. 1993; Hentze et al. 1999; Mendell et al. 2001). The gene responsible for Beta thalassemia, the beta-globin gene, provides the classic example of a primary splicing defect. Here, a single base substitution in the small intervening sequence (IVS1) of the beta-globin mRNA precursor results in alternative splicing and retention of the 3′ terminal 19 bases of IVS1 ahead of exon 2 (Fukumaki et al. 1982). There are numerous examples of endocrine disease resulting from primary splicing defects.

Isolated growth hormone deficiency type II is an autosomal dominant disease characterised by the failure to generate and traffic full length growth hormone (22 kDa). Several mutation affecting the splice sites of exon 3 or key ESEs result in

exon 3 being skipped and the production of a defective 17.5kDa growth hormone transcript. Exon skipping appears to be mediated by enhanced SC35 binding to local ESSs (Solis et al. 2008).

Laron's syndrome, or inherited growth hormone insensitivity is a syndrome characterised by mutations causing disruption of the growth hormone receptor. A point mutation has been identified which strengthens a 5' splice site for an intronic pseudoexon. One hundred and eight nucleotides between exon 6 and 7 are included in the growth hormone receptor transcript resulting in the addition of a 36 amino acid sequence which renders the receptor defective (Metherell et al. 2001).

Mutations of the genes encoding the succinate dehydrogenase subunit are known to predispose to the familial paraganglioma syndromes associated with catecholamine excess. We have demonstrated that a novel mutation in a subject with multiple paragangliomata (c. 169 + 1 G>A) results in disruption of the 5' splice site of exon 2 which is thus skipped, leaving a defective transcript including exon 1,3 and 4 (Srirangalingam et al. 2010).

Medium-chain acylCoA dehydrogenase (MCAD) deficiency is an example of a disease in which disruption of a splicing regulator element leads to an alteration in splicing. Patients with a missense mutation in exon 5 of the medium-chain acylCoA dehydrogenase (MCAD) are unable to generate a functional protein leading to MCAD deficiency and a resulting syndrome of hypoglycaemia, hyperammonaemia and increased risk of sudden death due to an inability to mobilise fatty acid for oxidation and energy production. The missense mutation inactivates an exonic splicing enhancer which is required to select a sub-optimal 3' splice site for exon 5 (Nielsen et al. 2007).

Thyrotropin-secreting pituitary adenomas (TSHomas) are rare pituitary tumours characterised by a raised thyroid stimulating hormone (TSH) levels despite

elevated thyroid hormone levels due to the failure of negative feedback. Some TSHomas contain an aberrantly spliced variant of the thyroid hormone receptor beta 2 (TRβ2spl) lacking 40 amino acids of the ligand binding domain (Ando et al. 2001). Thus T3, the ligand, is unable to bind the receptor which would normally result in down regulation of TSH secretion.

Splicing mutations are known to be involved in the disease process of several other endocrine tumours including insulinoma, thyroid and prostate carcinoma (Minn et al. 2004; Narla et al. 2005; Baitei et al. 2009).

1.2.17.2 Secondary splicing defects.

Secondary splicing defects may result from mutations in genes coding trans-acting components of the splicing machinery. Unlike cis-acting mutations causing primary splicing defects in one gene, mutations affecting components of the splicing machinery have the potential to affect many other genes. Diseases associated with such defects include Prader-Willi syndrome and retinitis pigmentosa (Kishore et al. 2006; Deery et al. 2002).

Prader-Willi syndrome results from the loss of the maternally imprinted region of chromosome 15q11-q13 via a number of mechanisms (deletions of the paternal region accounting for the majority of cases). Though many genes have been identified in this region e.g. SNURF-SNRPN, NDN, MKRN3, MAGEL2, PWRN1, PWRN2, IPW, PAR-1, PAR-4, PAR-5, C15orf2, it remains unclear which ones are responsible for the phenotypic features of disease (Butler 2011). One of these gene loci, the SNURF–SNRPN locus encodes the small nuclear RNA, C/D box snoRNA-HBII-52. A secondary splicing defect results from the loss of expression of HBII-52 (Kishore et al. 2006). HBII-52 binds a silencer element in exon Vb of the human serotonin receptor 2C mRNA leading to exon exclusion. The lack of HBII-52 in

affected subjects leads to alternative splicing and an abnormal serotonin receptor. Bioinformatic and experimental studies have identified five further pre-mRNAs (DPM2, TAF1, RALGPS1, PBRM1 and CRHR1) that appear to be alternatively spliced by the mouse homolog of the human snoRNA HBII-52, MBII-52 (Kishore et al. 2006).

1.2.17.3 Tertiary splicing defects.

Tertiary splicing defects cause disease as a result of abnormalities in splicing via mechanisms other than primary or secondary splicing defects.

Myotonic dystrophy, a microsatellite-expansion disorder, characterised by myotonia and insulin resistance provides an example of a tertiary splicing defect. This occurs, in part, as the result of sequestration of the splicing factor muscleblind like protein 1 (MBNL1) in nuclear foci made up of a CTG trinucleotide expansion in the 3′ untranslated region (3′ UTR) of the dystrophia myotonica-protein kinase (DMPK). These expanded RNA products localise as nuclear foci. The lower levels of the MBNL1 splicing factor result in a misregulation of alternative splicing for a subset of genes leading to the clinic phenotype of disease (Kuyumcu-Martinez et al. 2007).

1.3 Modification of Splicing.

1.3.1 Chemical and hormonal modification of splicing.

Splicing may be modulated both chemically and hormonally. Many chemical compounds have been identified which may modulate splicing choices and may provide potential therapeutic options (Bracco et al. 2003; Soret et al. 2006). These compounds are thought to work via several different mechanisms including alterations in; the steady state of splicing transcripts; the phosphorylation states of

SR proteins; the intracellular distribution of splicing factors and protein–protein interactions (Xiao et al. 1997; Graveley 2000). Several large scale screens to identify inhibitors for various splicing factors have been undertaken (Blanchette et al. 2005; Soret et al. 2006). For example, several indolozole derivative compounds have been identified which inhibit the kinase activity of topoisomerase I (topo I), a compound which is known to phosphorylate SR proteins (Soret et al. 2006).

Many examples of hormonally mediated splicing events have now been documented (Akker et al. 2001). These include the regulation of the IR exon 11 splicing by dexamethasone, insulin and glucose and the regulation of IGF-1 splicing by growth hormone (Chew et al. 1995).

1.3.2 Antisense strategies.

The unifying element of antisense strategies is the specific binding of an oligonucleotide sequence by Watson and Crick base-pairing to a given pre-mRNA sequence. The chemistry of the oligonucleotides backbone and additional modifications may vary and can influence efficacy and function. Antisense strategies broadly work in one of two ways. Base-pairing results in either the RNA being degraded or modified. RNA degradation may be caused by one of several mechanisms including degradation via RNaseH, RNaseP or RNA interference via siRNA and miRNA. Antisense strategies to modify RNA may occur by a number of different mechanisms:

1. Blockade of cis-acting elements e.g. splice sites.
2. Recruitment of trans-splicing factors e.g. TOES and ESSENCE.
3. Replacement of spliced mRNA via a trans-splicing reaction e.g. SMaRT.

1.3.2.1 Antisense strategies – RNA degradation

The first ASO approved by the FDA was Fomivirsen (Vitrvene) in 1998, a phosphorothioate oligonucleotide against Cytomegalovirus (CMV). This was achieved by targeting the major immediate early region 2 (IE2) mRNA of the human CMV for degradation via RNaseH. IE2 protein is essential for virus survival. RNaseH is an endonuclease which degrades RNA within RNA-DNA hybrids by hydrolysis of phosphodiester bonds. Administration, via intraocular injection was indicated for the treatment of CMV retinitis in subjects with acquired immunodeficiency syndrome (AIDS).

Gapmers were second generation ASOs which consist of a central chain of nucleosides, typically a phosphothorothiate backbone flanked by nucleosides with modifications e.g. 2'-*O*-methoxyethyl (2'-MOE) and 2'-*O*-methyl (2'-OMe) which rendered them resistant to degradation by exonucleases but still a target for degradation by RNaseH. This antisense technology is being applied for therapeutic benefit. A large number of oligonucleotides are now in phase I and II trials and several compounds have reached phase III studies. Isis pharmaceuticals have developed Mipomersen targeting APOB 100 and OGX-011 targeting Clusterin. Mipomersen is a gapamer with a phosphorothioate oligonucleotide and 2'-O-methoxyethyl modified ribose sugars at either end. Mipomersen targets the full length APOB 100, the lipoprotein component of the atherogenic LDL responsible for familial hypercholesterolaemia with a predisposition to cardiovascular disease.
OGX-011 is another gapmer targeting clusterin, a secreted protein involved in resistance to apoptosis and anti-cancer agents. It has been used as adjunctive therapy in association with docetaxel for prostate, non-small cell lung and breast carcinoma and improved median survival has been noted in phase II trials with the

addition of the drug (Laskin et al.; Saad et al.; Chia et al. 2009).

Other degradative pathways such as that via RNaseP have been employed to target mRNA encoding a chloramphenicol-resistance protein in E. Coli using peptide-conjugated phosphorodiamidate morpholino oligomers (PPMOs) (Wesolowski et al. 2011). RNA interference, described in section 1.2.8.5, has been extensively exploited experimentally using siRNA to knockdown proteins but may also provide potential therapeutic options in the future.

1.3.2.2 Antisense strategies: Blockade of cis-acting elements.

ASOs may be designed to block key cis-acting elements, in a gene specific manner, which are critical to the splicing process. Importantly the ASO chemistry must be modified in order to prevent RNaseH mediated degradation. The use of such ASOs may result in either exon inclusion or exclusion. ASOs targets include constitutive or cryptic splice sites and enhancer (ESE, ISE) or silencer (ISS, ESS) sites. The mechanisms by which the targeting ASO cause the desired effect vary. Commonly, targeting ASOs causes exon skipping or inclusion which alters the resulting protein but other mechanisms also exist.

A recent phase II study using the antisense oligonucleotide PRO051, administered subcutaneously in subjects with muscular dystrophy, was shown to result in new dystrophin expression (which is lacking in affected subjects) of up to approximately 16% of normal expression levels in the tibilialis anterior muscle (Goemans et al. 2011). PRO051 is a 2′-O-methyl–substituted ribose molecule with phosphorothioate internucleotide linkages designed to bind to exon 51 and cause exon skipping. This is mediated by the conversion of an out-of-frame transcript to

an in-frame shortened transcript which is nevertheless functional. This was demonstrated to be clinically relevant as exertional capabilities increased with a short period of therapy.

Spinal muscular atrophy (SMA) is caused by a defective survival of motor neuron (SMN1) gene resulting in spinal neurone degeneration and consequent muscle wasting. The paralogous gene, *SMN2*, is unable to replace the defective *SMN1* as a C to T change at position +6 in exon 7, results in the loss of a SF2/ASF enhancer site and skipping of exon 7. An ASO targeting an ISS flanking exon 7 of the SMN2 gene resulted in exon inclusion thus partially restoring some SMN protein expression (Hua et al. 2007). This ASO (ASO-10-27) has been delivered systemically and intracerebroventricularly in mouse models of spinal muscular atrophy resulting in prolonged survival (Hua et al. 2011).

Other targets with translational potential include skipping of a pseudoexon in the β-globin gene using a PPMO targeting an intron 2 mutation (IVS2-654) responsible for β-thalassemia which results in low haemoglobin levels. This was combined with siRNA mediated degradation targeting the alpha globin subunit, levels of which tend to be elevated in β-thalassemia, to increase haemoglobin levels in a mouse model (Xie et al. 2011).

Another mechanism appears to mediate ASOs targeting the CUG expansion repeats in myotonic dystrophy. The ASOs bound to the CUG repeats appear to displace the binding of muscleblind like protein 1 (MBNL1), a splicing factor, the lack of which contributes to the mis-splicing of several genes which are responsible for the phenotype seen in subjects with myotonic dystrophy e.g. muscle-specific chloride channel protein 1. Restoration of splicing of the muscle-specific chloride channel protein 1 reduced myotonia in mice injected with a 25-mer

phosphorodiamidate morpholino oligomers (PMO) targeting the CUG repeats (Wheeler et al. 2009).

1.3.2.3 Antisense strategies: Recruitment of trans-splicing factors.

Tailed oligonucleotide enhancement of splicing (TOES) compounds consist of an ASO annealing base targeted to a designated site with an attached RNA sequence that mimics an ESE (Skordis et al. 2003; Ghigna et al. 2010). The purpose of the tail is to provide a trans-acting factor capable of recruiting either an enhancing or inhibitory splicing factor. Additional enhancements e.g. a U7 stem loop have been added in some cases to enhance the effectiveness of the compounds (Baughan et al. 2009; Meyer et al. 2009). These compounds have been extensively tested in exon 7 of the SMN2 pre-mRNA, involved in SMA both in vitro, in vivo in cultured fibroblasts from subjects with SMA, parenterally (peripherally and centrally) and transgenically in mouse models (Skordis et al. 2003; Meyer et al. 2009; Baughan et al. 2009). Increasing amounts of the ASO hybrid compounds increased the proportion of the exon 7 included in the SMN2 mRNA transcripts.

Exon-specific splicing enhancement by small chimeric effectors (ESSENCE) is a technique that has successfully been used in vitro to restore exon 18 inclusion in BRCA1, exon 7 inclusion in SMN2 (Cartegni et al. 2003) and in vivo to enhance splicing of the pro-apoptotic bcl-xS isoform in preference to the anti-apoptotic bcl-xL isoforms (Wilusz et al. 2005). This technique employs chimeric molecules incorporating an antisense PNA oligonucleotide targeting an exon, linked to a peptide tail containing ten RS (arginine/serine) repeats. This RS domain tail simulates the function of an SR splicing factor thus facilitating the inclusion of the targeted exon.

1.3.2.4 Antisense strategies: Replacement of mRNA via trans-splicing reaction.

Spliceosome-mediated RNA trans-splicing (SMaRT) creates a chimeric mRNA through a trans-splicing reaction mediated by the spliceosome between the 5′ splice site of an endogenous target pre-mRNA and the 3′ splice site of an exogenously delivered trans-splicing RNA molecule (Puttaraju et al. 1999). These compounds consist of a binding domain to an intronic region of the targeted gene and a trans-splicing domain containing coding sequence corresponding to the new 3′ sequence (Puttaraju et al. 1999). This technology has been used in several scenarios, notably the functional repair of the CFTRΔF508 mutation of the cystic fibrosis transmembrane conductance regulator in human airway epithelial cells (Liu et al. 2002) and the phenotypic correction of haemophilia A in mice (Mansfield et al. 2004).

Fronto-temporal dementia with parkinsonism linked to chromosome 17 (FTDP-17) is caused by mutations in the gene encoding the microtubule-associated protein, tau. A 34% efficiency in converting exon 10- tau RNA to exon 10+ RNA using this technology has been demonstrated in vivo (Rodriguez-Martin et al. 2005).

Mechanisms to enhance efficacy are becoming increasingly sophisticated. Shababi et al. report the use of an enhanced trans-splicing systems to rescue the disease phenotype in a mouse model of spinal muscular atrophy. A single plasmid vector system was constructed which coded a trans-splicing RNA (tsRNA), an ASO designed to block a downstream endogenous splice site and the cDNA for the neuroprotectant, IGF-1 (Shababi et al. 2011). Intracerebroventricular injection of this vector into an SMN mouse model resulted in increase levels of SMN in the central nervous system associated with significant improvements in weight gain and survival compared to untreated mice.

1.3.3 Antisense oligonucleotides chemistry.

Over time several modifications to the basic chemistry of the ASOs have been made to improve their delivery, stability and affinity. The transfection efficiency for an ASO with a ribose-phosphate backbone may be limited by its negative charge. The first generation ASOs were phosphorothioate (PS) modified in which a sulphur molecule replaces one of the non-bridging oxygen atoms. Phosphorothioate modified ASOs have an increased resistance to nuclease degradation, increased bioavailability and stimulate RNaseH mediated degradation when duplexed with a target RNA (Eckstein 2000).

Second generation modification involved the development of 2′-alkyl modifications of the ribose ring including 2′-O-methyl (2′-OMe) and 2′-*O*- Methoxyethyl (2′-MOE) modifications. These modifications further increase binding affinity to the target RNA and increased nuclease resistance. Importantly, these second generation modifications did not support RNaseH mediated degradation and were therefore useful when the remit was other than transcript degradation e.g. increasing splicing of an exon (Altmann et al. 1996).

Third generation modifications introduce further changes to the chemistry of the backbone such as peptide nucleic acids (PNA), locked nucleic acid (LNA) and phosphorodiamidate morpholino oligomer (PMO). PNA oligonucleotides replace the phosphodiester backbone with a pseudopeptide polymer (*N*-(2-aminoethyl)glycine). Bases were attached to the backbone via methylene carbonyl links (Egholm et al. 1993). The neutral charge of the PNA backbone increased cell permeability and stability. PNA molecules are resistant to degradation by RNaseH mediated degradation (Egholm et al. 1993). Locked nucleic acid (LNA) contains a 2′-*O*,4’-*C*-methylene bridge in the ribose ring which confers high stability and increased affinity to RNA. LNA binding does not stimulate RNAseH mediated degradation

(Vester et al. 2004; Kurreck et al. 2002). PMOs are generated by replacing the ribose ring with a morpholino ring and a phosphorodiamidate in place of the phosphodiester bond. Resistance to RNaseH and nuclease degradation is increased however limited cell penetration remains an issue (Amantana et al. 2005).

1.4 Insulin receptor.

1.4.1 Insulin receptor: introduction.

The insulin receptor (IR) cDNA (OMIM – 147670) is encoded by a genomic region located on chromosome 19p which is comprised of 22 exons (figure 1.4.1.A). The IR is synthesised as a ~180 kDa pro-receptor which is cleaved into α–β monomers, linked by disulphide bonds. The IR monomer homodimerises with itself to form the IR, or with IGF-I receptor to form a heterodimer (Soos et al. 1990). The α subunit is the extra-cellular domain that binds insulin. The ß subunit contains a short extracellular domain, a transmembrane domain and cytoplasmic domain. The cytoplasmic domain has an intrinsic tyrosine kinase activity (Ebina et al. 1985; Ullrich et al. 1985).

1.4.2 Insulin receptor and signalling cascade.

The binding of insulin to the α-subunit of IR leads to a conformational change resulting in the autophosphorylation of the tyrosine residues of the β-subunit (Frattali et al. 1992) (figure 1.4.2.A). The intrinsic tyrosine kinase activity of the β-subunit phosphorylates a number of docking proteins (IR substrate (IRS)) via a phospho-tyrosine binding module (White 1997; Giovannone et al. 2000). The IRS proteins dock with Src Homology 2 (SH2) domain proteins, one being the p85 regulatory subunit of phosphatidylinositol 3-kinase (PI3K). Movement of the catalytic p110α-subunit to the plasma membrane and phosphorylation of PI(4,5)-biphosphate (PIP_2) to PI(3,4,5)-triphosphate (PIP_3) follows (Fruman et al. 1998). PIP_3 activates the 3′ phosphoinositide-dependant kinase-1 (PDK-1) and protein kinase B (PKB or AKT). Substrates for PDK-1 include PKB and atypical forms of Protein Kinase C (PKC) (Kotani et al. 1998). These substrates mediate several

cellular metabolic changes (figure 1.4.2.A) including the effects of glucose transport via GLUT4.

Figure 1.4.1.A

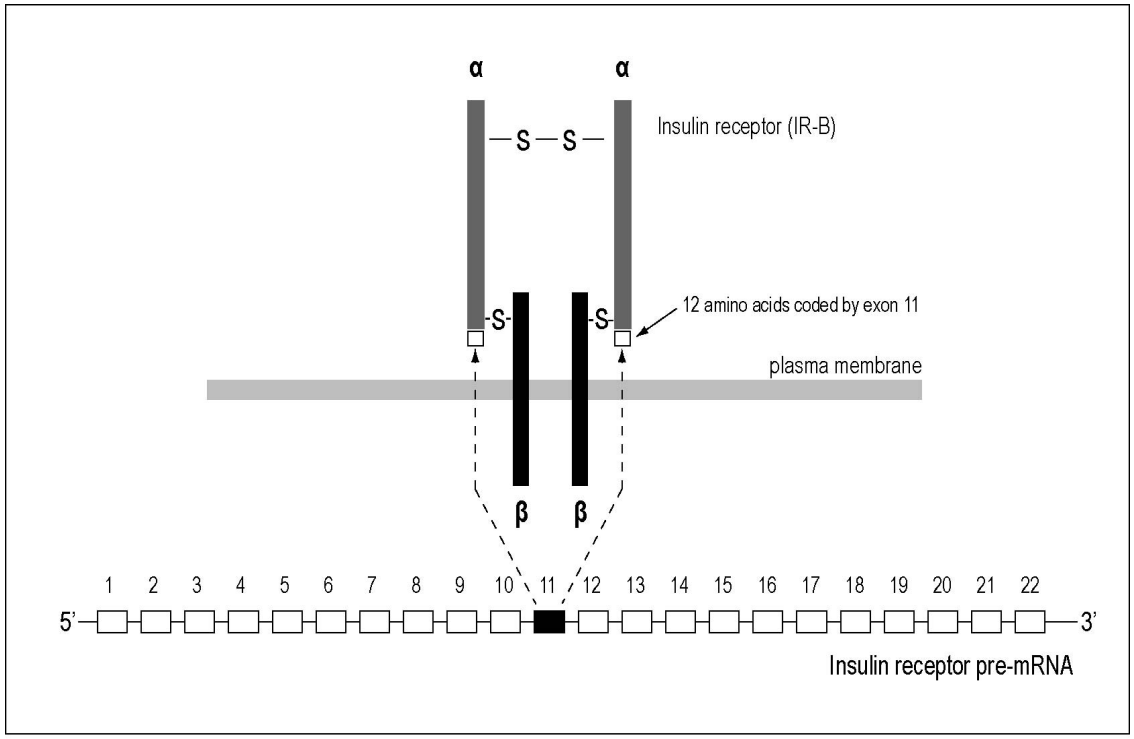


Figure 1.4.1.A The insulin receptor. The IR gene is located on chromosome 19 and consists of 22 exons. The IR is composed of a pair of α (extracellular) and β (transmembrane) subunits linked together by disulphide bonds. The 2 isoforms of the IR differ by the absence or presence of 12 amino acids at the carboxy terminus of the α subunit corresponding to the alternative splicing of exon 11. The insulin receptor A isoform (IR-A) excludes exon 11 and the insulin receptor B isoform (IR-B) includes exon 11 (shown in figure).

Figure 1.4.2.A

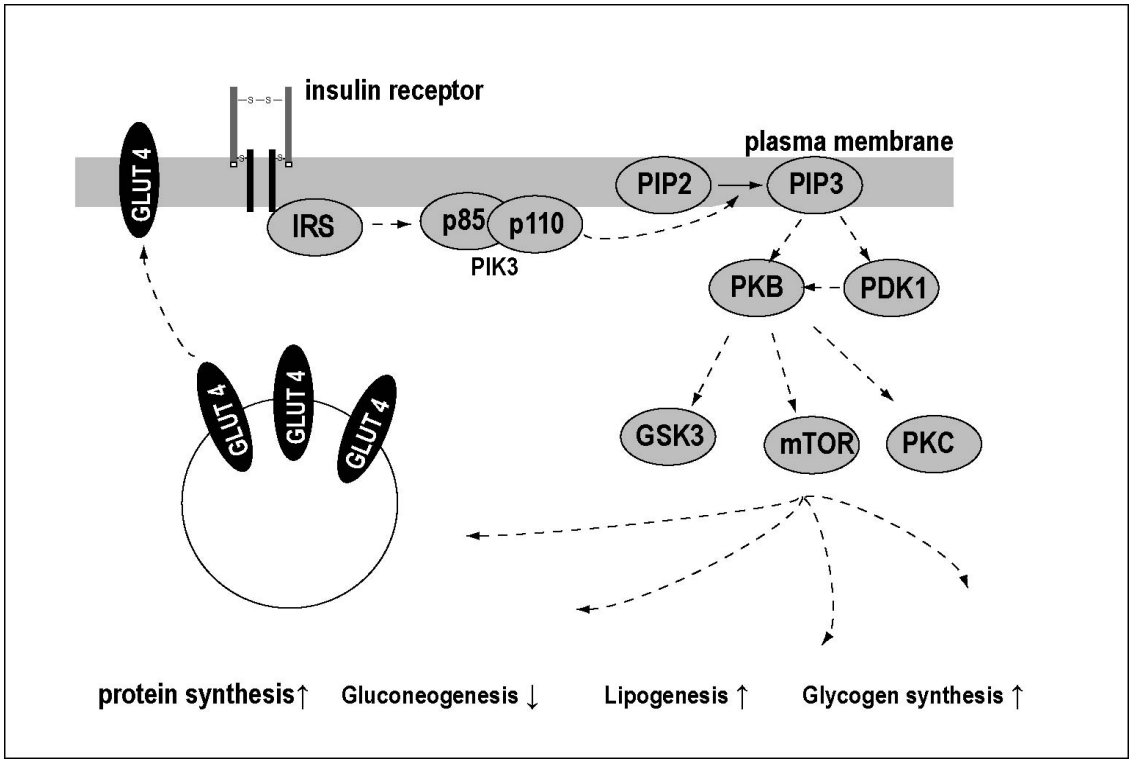


Figure 1.4.2.A Insulin signalling. Schematic diagram showing a portion of the cellular plasma membrane and downstream signalling via the IR. Insulin binds the IR and leads to a signalling cascade culminating in the translocation of GLUT4 from the intracellular pool to the plasma membrane to allow cellular glucose influx. Multiple other pathways and cellular effects are involved in insulin signalling but are not shown here for clarity. IRS - insulin receptor substrate; PI3K - phosphatidylinositol-3-kinase; PIP3 - phosphatidyl inositol-3,4,5-trisphophate; PIP2 - phosphatidylinositol-4,5-bisphosphate; PDK1 - PIP3-dependent protein kinase; PKB - protein kinase B (AKT); GSK3 - glycogen synthase kinase-3; mTOR - mammalian target of rapamycin; PKC - protein kinase C.

1.4.3 Role of insulin.

Insulin has wide-spread effects on cellular metabolism, growth, differentiation and survival. Insulin causes cellular glucose uptake, glycogen synthesis, lipogenesis and anabolic effects via protein synthesis. Indeed the importance of insulin and its effects via its receptor can be seen in mouse knockout studies. Disruption of the IR gene in mice pancreatic β cells results in impaired glucose secretion and glucose intolerance (Kulkarni et al. 1999). A more generalised knockout of the IR results in a 10% reduction in birth weight in mice which die after a few days due to keto-acidosis (Accili et al. 1996), suggesting a pivotal role of the IR in growth and glucose metabolism. Further tissue-specific knock-outs suggest roles for the IR in neovascularisation (Kondo et al. 2003), adipogenesis (Entingh et al. 2003), glucose disposal in muscle and adipose tissues (Lauro et al. 1998), and hepatic gluconeogensis (Fisher et al. 2003).

1.4.4 Alternative splicing of IR.

The elucidation of the structure of the IR revealed the presence of two isoforms (Ebina et al. 1985; Ullrich et al. 1985). These isoforms differ by 12 amino acids (residues 717-728) located at the C-terminus of the α-subunit. Alternative splicing of the in-frame 36 nucleotide exon 11 is responsible for this. The IR-A isoform excludes exon 11 and the IR-B isoform includes exon (Seino et al. 1989).

At 36 nucleotides in length, the IR exon 11 is short in comparison to the average exon, the result being that exon 11 is alternatively spliced and the proportion of each isoforms varies on the basis of the species and tissue involved.

The preponderance of the isoforms varies between tissues, during development, and pathophysiologically. IR-A is found exclusively in the spleen, peripheral blood cells and brain whilst IR-B is markedly predominant in the liver. There was noted to

be marginally more IR-B compared to IR-A in muscle, adipocyte and fibroblasts (Moller et al. 1989; Seino et al. 1989; Mosthaf et al. 1990). In fetal tissue, however, IR-A is the predominant isoform in kidney, muscle, liver and fibroblasts (Frasca et al. 1999). IR-A also appears to play a role in tumourigenesis, being upregulated in several form of cancer (Frasca et al. 1999; Sciacca et al. 1999).

1.4.5 Splicing influences on the insulin receptor exon 11.

Alternative splicing decisions appear to also be hormonally mediated. It has been demonstrated that dexamethasone can induce a switch in splicing from IR-A to IR-B in cultured HepG2 cells (Kosaki et al. 1993; Norgren et al. 1994) as can triiodothyronine to a lesser degree (Kosaki et al. 1993). Insulin appeared to significantly increase the proportion of IR-B in a rat hepatoma cell line (Sell et al. 1994) as did hyperglycaemia in human islets and RIN beta-cells (Hribal et al. 2003) though these findings were not confirmed in all studies (Norgren et al. 1994).

Intronic and exonic sites involved in alternative splicing of exon 11 in the IR have been more specifically studied using minigene constructs inserted via plasmids in HepG2 cells (Kosaki et al. 1998). A minigene is a simplified laboratory version of a natural gene that lacks one or more of the gene's exons and introns, or portions of them (Cartegni et al. 2002). Minigenes may be cloned into expression vectors and expressed in cells for in vivo analysis of alternative splicing mechanisms. To be a useful representation of an endogenous gene the minigene musts at a minimum have all the necessary RNA element of a gene to allow equivalent splicing of the minigene.

A 48-nucleotide purine-rich sequence at the 5′ end of intron 10 enhances exon 11 inclusion whilst a 43-nucleotide sequence at the 3′ end of intron 10 upstream of the branch point sequence causes exon skipping (Kosaki et al. 1998). Introducing a

four-point mutation in the middle of the exon 11 caused this to be spliced constitutively whilst a 8 nucleotide deletion towards the 3′ end or insertion of three thymidine residues resulted in significant exon 11 exclusion. Previous work from our laboratory has shown that mutating sub-optimal splice sites flanking exon 11 to consensus sequences causes it to be constitutively spliced in in-vivo and in-vitro models (Webster et al. 2004). Further evidence comes from studies involving CUG-BP1, a regulator of pre-mRNA splicing. Deletion of a 110 nucleotide intronic sequence abolishes the ability of this factor to up-regulate the splicing of IR-A in fibroblasts from patients with myotonic dystrophy (Savkur et al. 2001).

During the initial stages of this work the splicing factors involved in the splicing of the IR exon 11 had not been identified. In 2009 Sen et al. published data which identified the key players in this process (Sen et al. 2009). These findings, which are clearly relevant to the current work are discussed in detail in Chapter 4 in relation to the findings of this study.

1.4.6 What are the differences between IR-A and IR-B?

Differing properties have been noted between the isoforms though there is conflicting data in this area. The IR-A isoform has a two-fold higher affinity for insulin than IR-B in transfected mammalian cells (Mosthaf et al. 1990; Yamaguchi et al. 1991; Yamaguchi et al. 1993) however there is an equivalent response to insulin binding in terms of metabolic (glucose uptake) and mitogenic consequences (thymidine incorporation) (Yamaguchi et al. 1991). IR-B has a higher degree of autophosphorylation and greater kinase activity (Kellerer et al. 1992; Kosaki et al. 1995) but this was not found in all studies (Yamaguchi et al. 1991). Some studies using transfected cells suggest IR-A shows higher rates of internalization, recycling, and insulin-induced down-regulation of the receptor (Vogt et al. 1991; Yamaguchi

et al. 1991; Yamaguchi et al. 1993) though again, this is not confirmed in all studies (McClain 1991).

The insulin-like growth factors also bind the IR, and there is evidence for differences in the way they interact with the IR isoforms. Both IGF-I and IGF-II binds with higher affinity to IR-A versus IR-B (Yamaguchi et al. 1993; Frasca et al. 1999), IGF II binds with significantly higher affinity than IGF-1 to IR-A (Denley et al. 2004). However IGF/IR hybrid receptor studies carried out recently suggest that there is little difference between the isoforms in terms of ligand binding and activation of the receptor (Benyoucef et al. 2007).

1.4.7 Clinical significance of IR alternative splicing.

As IR-A and IR-B appear to differ in their affinity and their response to insulin, several groups have addressed the possibility that the alternative splicing of IR may play a key role in the pathogenesis of insulin resistance in type II diabetes. The data is conflicting however. Several studies quantifying either the presence of alternative mRNA or protein isoforms of the IR in samples of adipose tissue and skeletal muscle suggest an increase of IR-B isoforms in human subjects with insulin resistance or type II diabetes, compared with non-diabetic subjects (Mosthaf et al. 1991; Sesti et al. 1991; Kellerer et al. 1993; Mosthaf et al. 1993; Norgren et al. 1993; Sesti et al. 1995). However, other studies showed no difference in the ratio of the isoforms between subjects (Benecke et al. 1992; Anderson et al. 1993; Hansen et al. 1993). It has been shown in spontaneously obese diabetic Rhesus monkeys that hyperinsulinaemia (early diabetes) rather than hypoinsulinaemia (late diabetes) is associated with increased levels of IR-A in skeletal muscle (Huang et al. 1994). In the human liver both at an mRNA and protein level several studies have shown a marked predominance of IR-B of between 75-90% in non-diabetic

individuals (Moller et al. 1989; Seino et al. 1989; Benecke et al. 1992; Norgren et al. 1994). However hepatoma cell lines in the rat (Sell et al. 1994) and humans (Kosaki et al. 1993) shows the reverse. In the liver of diabetic rhesus monkeys there appears to be an increase in the proportion of IR-A (Huang et al. 1996). Studies using ASOs which acutely downregulated IR by 95% in mice did not alter insulin action on glucose production suggesting the role the IR plays might be relatively minor (Buettner et al. 2005). Several reasons have been postulated to explain the discrepancy between studies including different species, cell lines, detection methods, PCR protocols, contamination of samples and the variable metabolic status of the diabetic subjects from hyperinsulinaemia to hypoinsulinaemia (Huang et al. 1996; Sesti et al. 2001).

Both myotonic dystrophy type 1 and type 2 are associated with insulin resistance (Savkur et al. 2001; Savkur et al. 2004). Insulin resistance appears to be very strongly correlated to a predominance of IR-A in this condition. Studies of skeletal muscle cells cultured from these subjects shows a predominance of IR-A (Savkur et al. 2001). Mis-splicing of other genes such as cardiac troponin T (Philips et al. 1998) and the muscle chloride channel (Charlet et al. 2002) underlie the other clinical features of myotonic dystrophy, i.e. the muscular dystrophy and myotonia respectively. Overactivity of the splicing proteins CUG-BP and hnRNP H and loss of the activity of proteins in the muscleblind-like family have been postulated to cause these specific splicing abnormalities (Philips et al. 1998; Savkur et al. 2001; Ho et al. 2004; Paul et al. 2006).

The visceral and temporal distribution of the IR isoforms and the effect these changes play in insulin resistance may be significantly more complex than assigning a simple role for each isoform (Guiraud-Dogan et al. 2007). A study examining the distribution of the IR isoforms in transgenic mice containing the

human CUG-containing DMPK transcripts in various insulin sensitive tissues found that the isoform proportion varied in different tissues, at different times and associated with variable responses to glucose. Indeed the complexity extended to differences within different muscle groups at different times. (Guiraud-Dogan et al. 2007). This underlines the difficultly in matching the molecular changes with the metabolic consequences.

The IR may have a role in mitogenesis and tumour development. It has been shown that tumour cells show a different proportion of the IR isoforms compared with normal cells (Frasca et al. 1999; Sciacca et al. 1999). IR-A appears to be upregulated in several cancers including colonic (Frasca et al. 1999), breast (Sciacca et al. 1999), ovarian (Kalli et al. 2002) and thyroid cancer (Vella et al. 2002). In hepatoblastoma cells IR-B has been shown to be down-regulated (von Horn et al. 2001). Hybrid receptors of IGF-1R and IR are found in increased levels in thyroid (Belfiore et al. 1999) and breast cancer (Pandini et al. 1999). IGF-II has predominant mitogenic effects and only binds to IR-A, not IR-B (Frasca et al. 1999). It leads to cell proliferation in breast cancer cells (Sciacca et al. 1999) and resistance to apoptosis in leiomyosarcoma cells (Sciacca et al. 2002). Studies show that both that IGF-II and IR-A levels increase as thyroid cancer progresses suggesting a possible autocrine stimulatory loop between the two (Vella et al. 2002).

1.4.8 Insulin receptor: summary.

In summary it appears the relative proportions of the IR isoforms vary depending on tissue, species and underlying metabolic status. There is a significant body of conflicting data with regard to the relationship between insulin resistance and the proportion of IR isoforms though some explanations for this have been offered. It is

unclear whether these differences are a cause of, a result of, or merely an epiphenomenon of insulin resistance. The role of the IR-A isoform in insulin resistance in myotonic dystrophy and cellular proliferation in tumourigenesis appears more robust however. Clarifying differences between the IR isoforms has been hindered by experiments with methodological weakness. Many experiments have used isoform cDNAs over expressed in cell lines which do not naturally contain IR. In studies which use cell lines which naturally contain IR, the switch in alternative splicing is induced with dexamethasone and triiodothyronine, a method which is not specific and which is likely to alter many other metabolic and proliferative pathways.

1.5 Apolipoprotein B.

1.5.1 Apolipoprotein B: introduction.

Apolipoprotein B (APOB - OMIM 107730) is the main apolipoprotein in both chylomicrons and low density lipoproteins (LDL) but is also a component of VLDL, IDL and Lp(a) particles (Chan 1992). The human APOB gene is located on chromosome 2 and consists of 29 constitutively spliced exons. As with the IR there are 2 principal isoforms derived from the APOB gene, though their means of generation are distinct. The APOB100 isoform represents the full transcript while APOB48 is a C-terminal truncated isoform (literally 48% of APOB100). APOB100 is synthesized in the liver, is required for the assembly of lipoproteins and is the ligand for the LDL receptor. Elevated levels of APOB100 are associated with raised levels of LDL and Lp(a) and a concomitant rise in cardiovascular risk via accelerated artherogenesis (Jialal 1998). APOB48 is present only in the intestine and is involved in chylomicron assembly and transport of intestinal fat. It lacks the C-terminal domain which is required for lipoprotein binding to the LDL receptor. As a result chylomicron clearance occurs via Apolipoprotein E (APOE) and a chylomicron remnant receptor instead of the LDL receptor (Chester et al. 2000). The truncated APOB48 isoform is generated within the intestine following RNA editing via a complex known as the editosome. Editing occurs in exon 26 at position 6666 of the mRNA converting a C to a U, generating a PTC to produce the C-terminal truncated isoform (Chester et al. 2000). This isoform is then able to join the chylomicron and allows intestinal fat to be absorbed. Clinical interest in APOB exon 26 lies in the fact that preferential conversion to APOB48 in preference to APOB100 may lead to beneficial effects in the lipid profile which could be exploited therapeutically.

Figure 1.5.1.A

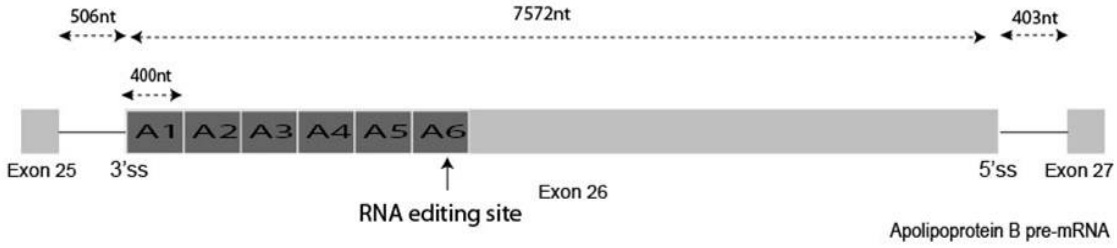


Figure 1.5.1.A. Apolipoprotein B exon 26. Schematic diagram shows APOB pre-mRNA including exon 25,26 and exon 27 (not to scale). Intron and exon lengths given. The 5' end of APOB exon 26 was divided into 6 regions 400 nucleotides wide, designated A1, A2.....A6. These regions were utilised experimentally to identify ESSs (see section 1.5.4). Regions A6 contains the RNA editing site at position 6666 in the mRNA marked with an arrow.

1.5.2 Apolipoprotein B: a therapeutic target.

Reductions in the level of APOB should lead to a reduction in LDL cholesterol and hence cardiovascular risk. Given this, APOB continues to be an important target for therapeutic manipulation (Akdim et al. 2007). Several approaches have been employed. Stimulating RNA editing by overexpression of the catalytic subunit of the editosome, APOBEC-1, increases the conversion of atherogenic APOB100 to APOB48. This has been associated with a reduction of LDL levels in transgenic mouse models (Yamanaka et al. 1995). Further approaches include ASO mediated degradation of APOB mRNA (Soutschek et al. 2004; Crooke et al. 2005; Zimmermann et al. 2006). There remains concern that knockdown of APOB100 would result in lower levels of APOB48 and lead to intestinal fat malabsorption. Evidence for the effect of reducing APOB synthesis comes from several sources. Patients with familial hypobetalipoproteinemia (FHBL) provide some insight into the effect of reducing APOB. Heterzygote subjects have a 50% reduction in the levels of APOB and appear to have a reduce risk of developing atherosclerosis (Tarugi et

al. 2007) while homozygotes, who have very low levels of APOB suffer with fat and fat soluble vitamin malabsorption, progressive neurologic degenerative disease, retinitis pigmentosa, acanthocytosis (spiculated red blood cells) and hepatic steatosis. Thus it is clear that that the aim of therapy would be a reduction of APOB rather than complete elimination. A more precise strategy would be to reduce the APOB 100 isoform without reducing the APOB 48 isoform thus reducing LDL cholesterol without affecting intestinal absorption.

Studies of the APOB synthesis inhibitor Mipomersen (Isis pharmaceuticals) a 20-mer antisense chimeric phosphorothiate oligonucleotide targeting APOB reveal than 10 to 15% of subjects developed a transaminitis in phase II and III studies (Akdim et al. 2010; Raal et al. 2010).

1.5.3 APOB exon 26 - long constitutive exon

Long exons such as APOB exon 26 contain many pseudo splice sites. Given this, correctly orientated pseudo 3′ and 5′ splice-sites should trigger exon definition and erroneous splicing within the exon 26 sequence. The genuine splice sites may however be identified via a process of intron definition. This relies on the recognition of 5′ and 3′ splice site pairs over short introns of <100 nucleotides in Drosophila (Talerico et al. 1994). In human nuclear extracts and in HeLa cells, intron definition appears to operate below a threshold of 200-250 nucleotides (Fox-Walsh et al. 2005). Intron definition, therefore, cannot easily explain the splicing of APOB exon 26 as the upstream and downstream introns are too long at 506 and 403 nucleotides respectively.

1.5.4 Apolipoprotein B: previous work.

Previous work done in our laboratory suggests that APOB exon 26 contains many pseudo splice sites within it, some of which are predicted to be stronger that the constitutive splice sites. It was hypothesised that exon 26 contains multiple ESS which block these alternative splice sites from being recognised. Using an in vitro heterologous splicing reporter, DNA ligase III, substituting sequences (400 nucleotides long) of exon 26 into an ESS position in the reporter allowed the definition of areas which were rich in silencing elements. These areas were delineated further by using 25 nucleotide sequences within a given 400 nucleotide sequence. The activity of these sequences was further confirmed in the β-globin splicing reporter HβΔ6. There appeared to be multiple silencer regions heterogenously spread through-out the exon with particular areas e.g. the central 4800 nucleotides being rich with these silencer elements.

1.6 Why study the IR exon 11 and APOB exon 26?

The IR exon 11 and APOB exon 26 provide examples of exons at the extremes of the size spectrum. The IR exon 11 and the APOB exon 26 are some of the shortest and longest spliced exons in the genome. They contrast in relation to the fact that the IR exon 11 is alternatively spliced while APOB exon 26 is constitutively spliced despite being significantly longer than the optimal length defined for efficient splicing. While the basic principles of splicing will hold their relative importance will vary considerably in the context of each model. Understanding how splicing is regulated as a result of these differences is of interest. In the case of the IR exon 11 - how does its small size impact on the ability to manipulate the splicing process? In the case of APOB exon 26 - how does the configuration of ESE and ESS within the pre-mRNA impact on the reproducibility of splicing in long exons?

1.7 Why use antisense oligonucleotides?

Antisense strategies provide an attract approach to modulating the splicing process for a number of reasons. Watson and Crick base-pairing allows the precise targeting of specific pre-mRNAs. ASOs provide a variety of strategies to modulate splicing which can be customised to each splicing scenario. They may be used to increase our understanding of the splicing process both as a normal molecular mechanism but also in disease as a pathological process. Antisense technology has already received approval for therapeutic use in some cases with a significant amount of drug development in the pipeline. As antisense chemistries and drug delivery systems refine such technology will become more applicable for therapeutic use.

1.8 Aims.

The ability to modify the splicing process provides significant opportunities to both study endocrine systems and manipulate them for therapeutic purposes. Previous studies utilise hormonal manipulation and minigene approaches which are cumbersome, non-specific and not easily translatable for therapeutic application. By contrast, the advantage of work with DNA and RNA is that Watson-Crick base-pairing allows the specific targeting of putatively important sequence elements by oligonucleotides with technology which is readily translatable.

The aim of the current study therefore is to use ASO technology to modify the splicing choices in two contrasting metabolic models of endocrine splicing, namely the short alternatively spliced IR exon 11 and the long constitutively spliced APOB exon 26.

More specifically, with regard to the alternatively spliced IR exon 11, the aim was to use ASOs with varying chemistries and modifications to cause exon 11 skipping and inclusion. This included the use of PNA-peptide hybrid ASOs to drive exon 11 inclusion.

Previous analysis has suggested the long constitutively spliced APOB exon 26 contains pseudoexons which remain unspliced in preference to the constitutive exon, as a result of the presence of multiple inhibitory exonic splicing silencer (ESS) elements along its length. The aim of the current study was to build on this work by confirming that ASOs, targeting some of these silencer elements can result in aberrant splicing of the normally constitutively spliced APOB exon 26 in an in vivo model.

Chapter 2

Material and methods

2.1 Antisense oligonucleotide design.

ASOs were designed complementary to the DNA sequence of the IR and the APOB gene as taken from the *Ensembl Genome Browser* (http://www.ensembl.org/index.html). Figure 2.1.A illustrates the base-pairing design for the 2′-O-methyl RNA ASO INSR ex1+2.

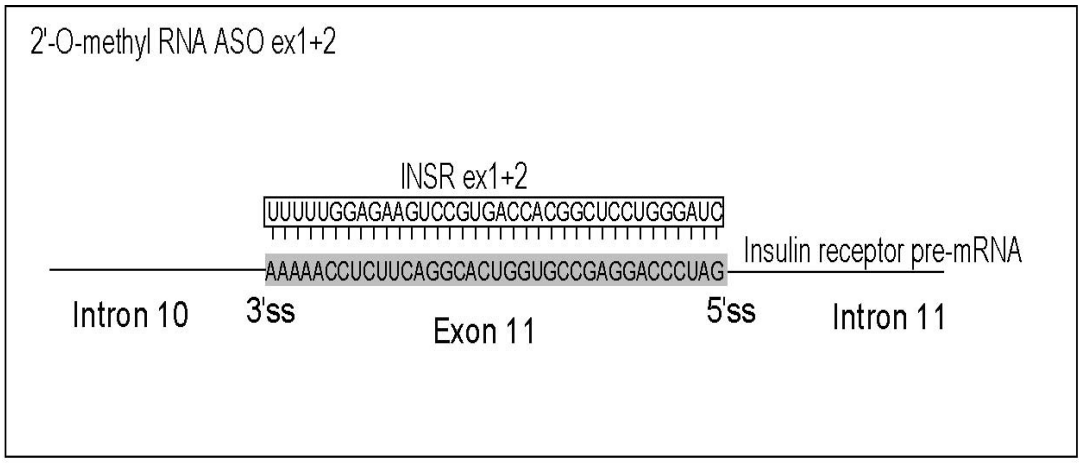


Figure 2.1.A ASO design: base-pairing with the IR exon 11.

2.1.1 ASOs targeting the IR exon 11.

2.1.1.1 2′-O-methyl RNA ASOs.

2′-0-methyl RNA ASOs were ordered from Dharmacon (Lafayette, CO, USA). Sites targeted for blocking intronic regions surrounding the IR exon 11 included the 3′ splice site, 5′ splice site and branch point sequence (see Chapter 3.1.1). These sites were identified and 18mer 2′-O-methyl RNA ASOs were designed. ASOs combining these sites were designed i.e. INSR 3B, INSR 5B, INSR 53. These were 36mers. Exonic ASOs were designed as 18mers (INSR ex1, INSR ex2, INSR ex3) and a combination ASO (INSR ex1+2) spanning the whole exon was designed (36mer). Sequences are listed in Chapter 6.1 - appendix. Two further ASOs were

designed and tested following initial experiments with the aforementioned ASOs. These included the intronic ASO INSR 6/7 and the exonic ASO INSR ex3.

2.1.1.2 Scramble ASOs.

Scramble ASO sequences of targeted ASOs to IR exon 11 were designed as controls (see Chapter 3.1.2). These scramble sequences had the same composition and number of oligonucleotides as the targeted ASOs. These ASOs were designed using a web-based tool (www.sirnawizard.com) to scramble sequences.

2.1.1.3 Immunofluorescent ASOs.

The effective ASO INSR ex1+2 and its scramble sequences INSR ex1+2scram had 5′ fluorescein tails added (see Chapter 3.1.4). These were designed to ensure ASO and scramble sequences localised appropriately and similarly.

2.1.1.4 TOES ASOs.

These modified ASOs (see Chapter 3.1.5) consisted of a 2′-O-methyl RNA ASO backbone (INSR ex1) with the addition of a 3′ or 5′ tail which mimicked an ESE (Chapter 6.1 – appendix for sequence) as used in another study (Skordis et al. 2003). This sequence was utilised because it has been shown to be efficacious but also because it attracted the splicing factor SF2/ASF which has been postulated to be relevant in the IR exon 11 model (Sen et al. 2009).

2.1.1.5 PNA ASOs.

The basic backbone used for PNA design was the repeating N-(2-aminoethyl)-glycine unit. Various modifications to the PNA backbone have been trialled to improve the functionality of the ASO with respect to DNA binding, sequence specificity and bioavailability. Modifications include those with acyclic structures such as substitution or addition to the C2 or C5 carbon or PNAs with cyclisation of the backbone either on the aminoethyl side or on the glycine side. (Corradini et al. 2011). The PNA backbones here were designed as three 12mers across the exon (INSR PNA ex4, INSR PNA ex5, INSR PNA ex6) (see Chapter 3.2.). A further PNA and its PNA scramble equivalent were designed to match the 2′-O-methyl RNA ASO INSR ex2. These were designated INSR PNA ex2 (18mer) and INSR PNA ex2scram (18mer). Two further 15mer PNA ASOs targeting the 5′ and 3′ end of exon 11, named INSR PNA ex1 and INSR PNA ex2 respectively were also designed. These two ASOs did not bind the 3 bases at the exon/intron boundaries and were used as the basis of the PNA-peptide hybrid ASOs.

2.1.1.6 PNA-peptide hybrid ASOs.

The PNA base was placed at both the 3′ and 5′ end of exon 11 in the equivalent INSR ex1 and INSR ex2 positions minus three nucleotides at the exon/intron junctions (see Chapter 3.2.3 and 3.2.4). It was assumed that an ASOs binding near or across the splice site may interfere with splice site recognition and inhibit exon splicing thus reducing the potential efficacy of compounds designed to enhance IR exon 11 inclusion. A glycine/glycine bridge then attached the RS (arginine/serine) repeat tail. The RS domain was repeated x10 in keeping with previous studies which showed an optimum response with this length of tail (Cartegni et al. 2003).

2.1.2 Apolipoprotein B exon 26 ASOs.

2′-O-methyl RNA ASOs were designed on the basis of previous work from our laboratory to identify ESS within APOB exon 26 (unpublished data). We identified the strongest ESS sequences (APOB exon 26 - 65-89(1), 2319-2343 (4)) based on 2 in-vitro reporter systems (DNA ligase III and the β-globin splicing reporter HβΔ6) and designed complementary sequences (Chapter 6.1 - appendix). We chose 2 adjacent ESSs close to the strongest ESS (APOB exon 26 - 97-121(2), 129-153(3)) in order to test the effects of blocking multiple adjacent silencer elements. Regions which had no apparent silencing activity were also targeted (APOB exon 26 - 93-217(5) and 2191-2215(6)) as controls (see Chapter 3.3).

2.2 Cell culture.

2.2.1 Cell lines.

The cell line used in this study was the HepG2 cell line (Human hepatoblastoma derived cell line). Cells used were between passage 2 and 23. HepG2 retain many liver specific functions including the production of apolipoproteins. (Pullinger et al. 1989). APOB, with a molecular mass of approximately 500kDa, is secreted from HEPG2 cells and the 22 KDa mRNA coding for APOB can be isolated from these cells, suggesting the full length APOB is produced (Thrift et al. 1986; Huang et al. 1985).

2.2.2 Maintenance of cells.

Materials:

- Minimum Essential Medium (MEM), (Invitrogen, Carlsbad, CA, USA).
- Foetal Bovine Serum (FBS), (Sigma, Dorset, UK)

Method:

The maintenance media consisted of MEM and 10% FBS. This was pre-warmed to 37ºC before use. Media was added to the flask/dish containing cells. Media was replaced every 48 hours. The passage number was recorded on each flask/dish.

2.2.3 Passaging cells.

Materials:

- Autoclaved 1x PBS (pH 7.4) (Sigma, Dorset, UK)
- Trypsin/EDTA (Tryple, Gibco/Invitrogen, Paisley, UK)
- Minimum Essential Medium (MEM) + 10 FBS%

Method:

Following 48 hours incubation of HepG2 cells in MEM/10% FBS, the petri dish was removed from the incubator and placed under the microscope to estimate cell confluency. If greater than 80% confluence, the cells were prepared for division. The media in the petri dish was aspirated. Cells were then washed in 8ml of PBS. PBS was aspirated and 1.5ml of trypsin/EDTA was added per 100mm petri dish. The petri dish was then placed in the incubator at 37ºC for 10 minutes to activate the trypsin. Following this, the dish, once removed from the incubator, was gently tapped to ensure the cells were displaced. Fresh MEM/10% FBS (8.5ml) of was added to the dish to neutralise the effects of trypsin. The cell suspension was placed in a 50ml Falcon tube. With a 200µl pipette tip inserted onto the tip of a 5ml pipette the cell suspension was drawn through the pipette 5-10 times in order to disperse the HepG2 cells. Then the cell suspension was divided into 2-3 petri dishes with the addition of further MEM to a final volume of 15ml/dish. The petri dish was then labelled with the passage number and placed in the incubator at 37ºC, 5% CO_2 .

2.2.4 Thawing of cells.

The cryotube containing 1ml of frozen cell suspension in FBS/DMSO was removed from the liquid nitrogen storage unit and thawed in a 37ºC incubator for approximately three minutes. The cell suspension was then added to a cell culture flask with fresh media at 37ºC and incubated.

2.2.5 Freezing down cells.

Materials:

- FBS (Sigma, Dorset, UK)

- DMSO (Sigma, Dorset, UK)

Method:

Cells were trypsinised as described above. The cell suspension was then spun at 1000rpm for three minutes. The supernatant was aspirated and the cell pellet was resuspended in 1ml FBS+10% DMSO. The cell suspension was placed in a cryotube and placed in an isopropranol freezing chamber in a -80ºC freezer to allow gradual freezing (1ºC/min). Once frozen the cells were transferred to the liquid nitrogen store within 24 hours. The passage number was recorded on the cryotube.

2.2.6 Counting cells.

Materials:

- Neuberg haemocytometer

Method:

Cells were counted in a Neuberg haemocytometer to determine cell concentration. The chamber was cleaned with 70% ethanol. A cover slip was then placed over the counting area. 25µl of trypsinised cells diluted in media (HepG2 cells from a 100mm Petri dish dissolved in a 10ml suspension) was injected beneath the cover slip. The $1mm^2$ counting grid was focussed on a microscope and the cells within it were counted. Cell concentration ($x10^4$/ml) was determined on the average of two counts per a given field view.

2.2.7 Reverse-transfection of ASOs into mammalian cell lines.

Materials:

- Lipofectamine 2000 (Gibco/Invitrogen, Paisley, UK)
- MEM/10% FBS (Gibco/Invitrogen, Paisley, UK)
- Opti-MEM I (Gibco/Invitrogen, Paisley, UK)

Method:

HepG2 cells were divided once 80% confluent as per the protocol (Chapter 2.2.3). Media was added to trypsinised cells. The cell concentration was determined with a Neuberg haemocytometer and diluted with further media to a concentration of 50 × 10^4 cells/ml. 1ml of cell suspension each was added to the required number of wells in a twelve well plate. One well was required for one transfection reaction. Lipofectamine 2000 and ASOs were gently mixed with relevant volumes of warmed Opti-MEM I and left to stand for five minutes. The lipofectamine 2000/Opti-MEM I (100µl) and ASO/Opti-MEM I (100µl) solutions were gently mixed to the relevant volume (200µl) and left to stand for twenty minutes. The lipofectamine 2000/ASO/Opti-MEM I solutions (200µl) were added to each well of cell suspension and gently mixed and placed in the incubator at 37ºC for 48 hours. Total volume per well was 1.2 ml.

2.3 RT-PCR.

2.3.1 Primers.

Primers for human IR exon 11 and APOB exon 26 are shown in Chapter 6.1 - appendix. Primers were ordered from Sigma (Sigma, Dorset, UK). Primer design was either based on primers reported in the literature specific to the sequence required or designed from sequences published in the DNA genomic database; ensembl (www.ensembl.org). Primers were designed using the program, Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer 3plus.cgi).

2.3.2. dNTP mix (10nM).

Materials:

- 100nM stock dATP, dTTP, dCTP, dGTP (Promega, Southampton, UK)

- RNase free water (Qiagen, Valencia, CA)

Methods:

A 10nM solution of dNTP (dATP, dTTP, dCTP, dGTP) was prepared from 100nM stock solutions of each deoxynucleotide according to the following mix:

	Volume
dATP (100nM)	20µl
dTTP (100nM)	20µl
dCTP (100nM)	20µl
dGTP (100nM)	20µl
dH20	120µl
Total volume (10nM dNTP)	200µl

2.3.2 RNA extraction.

Materials:

- Monoject Insulin syringe (Tyco Healthcare, Gosport, UK)
- RNeasy Micro kit (Qiagen, Valencia, CA)
- TurboDNase (Ambion, Huntingdon, UK)

Method:

RNA extraction was carried out in an RNA free environment using the RNeasy Micro kit according to the manufacturer's instructions as follows. After 48 hours incubation, transfection media was removed from each well and 350µl of RLT/β-mercaptoethanol (containing guanidine thiocyanate) solution was added where cell concentration was less than $5x10^6$ cells. The contents of each well was syringed five times through a 25G Monoject Insulin syringe and transferred to an eppendorf tube. 350µl of 70% ethanol was homogenised with the contents of each eppendorf

tube. The contents (700 µl) were transferred to a spin column mounted inside an eppendorf collection tube. The columns were spun in a microcentrifuge at 10,000rpm for 30 seconds. The flow through was discarded. 350µl of RW1 wash was added to each spin column. Spin columns were then spun at 10,000rpm for 30 seconds and the flow through was discarded. DNA digestion was undertaken by adding a 70µl solution, consisting of 69µl RDD buffer and 1µl TurboDNase to each spin column and left for fifteen minutes at room temperature. Following this, 350µl of RW1 wash was added to each spin column and centrifuged at 10,000rpm for 30 seconds and the flow through was discarded. The eppenderf collection tubes were replaced. 500µl of RPE washing buffer was added to each column and spun at 14,000rpm for one minute. The flow through was discarded. Then 500µl of 80% ethanol was pipette into each column and spun at 14,000rpm for one minute and the flow through was discarded. The eppendorf collection tubes were then changed and spun at 14,000rpm for five minutes. The spin columns were then placed in new eppendorf tubes. To elute the RNA, 14µl of RNAse free water was added to each column and spun at 14,000rpm for 1 minute. Eppendorf tubes were labelled and immediately stored at -80ºC.

2.3.3 RNA quantitation.

2.3.3.1 QuantIt Ribogreen quantification.

Materials:

- QuantIt Ribogreen kit (Gibco/Invitrogen, Paisley, UK)
- Wallac spectrofluometer

Method:

Total RNA was quantified using the QuantIt Ribogreen kit according to the manufacturer's instructions. TE solution x1 was prepared from TE stock (x20) using RNase free water. The standard RNA provided (100µg/ml) was serially diluted in TE from a stock solution at a concentration of 2µg/ml to give the following concentrations; 31.25ng/ml; 62.5ng/ml; 125ng/ml; 250ng/ml; 500ng/ml; 1000ng/ml. RNA test samples were diluted one in ten in TE solution. 100µl of RNA standard and test solutions were placed in a 96 well plate in triplicate. Blank wells in triplicate contained TE alone. A 100µl of a ribogreen solution (stock solution diluted 1 in 200 in TE) was placed in each well. The sample fluorescence of the RNA standards and test samples were measured using a spectrofluorometer. An absorbance at 260 nm (A260) in a cuvette with a 1 cm path length of 0.05 corresponds to 2 µg/mL RNA. The concentration of RNA was quantified against standard curves generated for standard RNA samples.

2.3.3.2. NanoDrop® ND-1000 UV-Vis Spectrophotometer.

Early experiments made use of the Ribogreen Kit for RNA quantification. In subsequent experiments the NanoDrop spectrophotometer was used. Several test samples were measured using both methods that confirmed equivalent quantification (Table 3.1.1.B).

1µl of total RNA was pipetted onto the fibre optic cable receiving platform and the second fibre optic cable arm brought into apposition with the receiving platform. The spectrophotometer (using a CCD array) then analyses absorbance of the sample following a flash pulse of xenon light. Data is transformed to give a concentration of RNA in ng/µl.

RNA purity was assessed by calculating the ratio of the absorbance of an RNA sample at two wavelengths (A260/A280). Adequately purified RNA was deemed to have a ratio of between 1.8 and 2.

2.3.4 Reverse transcription (RT).

Materials:

- 100nM stock dATP, dTTP, dCTP, dGTP (Promega, Southampton, UK)
- Transcriptor reverse transcriptase (Roche, Burgess Hill, UK)
- Protector RNase inhibitor (Roche, Burgess Hill, UK)
- RT buffer (x5) (Roche, Burgess Hill, UK)
- Oligo dT_{15} primer (0.5µg/µl) (Roche, Burgess Hill, UK)
- RNase free water (Qiagen, Valencia, CA)
- Total RNA sample

Method:

cDNA was synthesized from extracted RNA using RNase free materials. 2.5µg of RNA per reaction was used. As the concentrations of the harvested RNA varied between samples the following calculation was undertaken to correct for this and give the required volume of sample to give 2.5µg of RNA per reaction:

Desired weight in µg (in this case 2.5 µg) / Actual concentration in µg/µl

The RNA sample was pipetted into a sterile RNase-free microcentrifuge tube and RNase-free water was then added to a total volume of 6µl. 0.5µl of oligo dT_{15} primer was then added and gently mixed. The tube was heated for ten minutes at 65ºC to remove secondary structure within the template and subsequently at 4ºC to maintain this. A -RT control sample was also mixed in which the RNA sample was

replaced by 3.5µl RNase-free water. The following RT mix was prepared and 3.5µl was pipetted and gently mixed to each tube to give a final volume of 10µl.

RT mix (per reaction)

- RT buffer (x5) 2 µl
- dNTPs (10nM conc. dATP, dTTP, dCTP, dGTP) 1 µl
- Protector RNase inhibitor 0.25 µl
- Transcriptor reverse transcriptase 0.25 µl

Each RT sample was placed in the PCR machine and samples were subjected to the following programme:

	Temp		**Time/min:sec**
1.	65°c	for	10:00
2.	4°c	for	continuous
			add RT mixture
3.	50°c	for	60:00
4.	85°c	for	5:00
5.	End		

2.3.5 PCR reaction.

PCR reactions were carried out to amplify cDNA. FastTaq DNA polymerase was utilised for PCR reactions relating to IR exon 11. EXL DNA Polymerase, more suitable for the PCR of longer products, was used for the PCR of products from experiment involving APOB exon 26.

2.3.5.1 FastTaq DNA Polymerase PCR.

FastTaq DNA Polymerase was used in PCR for the IR exon 11.

Materials:

- 10nM concentration dATP, dTTP, dCTP, dGTP (Promega)
- FastTaq DNA polymerase, (Roche, Burgess Hill, UK)
- PCR buffer x10 ($MgCl_2$), (Roche, Burgess Hill, UK)
- Forward + reverse primer (concentration 50μM), (Sigma, Dorset, UK)
- [α-^{32}P] dCTP (Perkin-Elmer, Wellesley, USA)
- dH_20

Methods:

The volume of [α-^{32}P] dCTP added per reaction was adjusted according to α-^{32}P activity. Specific primer pairs were used for a given reaction (Chapter 6.2 - appendix). The typical reaction mixture was as follows:

PCR buffer (x10)	1 μl
dNTPs (10nM conc. dATP, dTTP, dCTP, dGTP)	0.2 μl
Forward primer	0.2 μl
Reverse primer	0.2 μl
FastTaq DNA polymerase	0.1 μl
0.05 μl (0.5μ Ci) [α-^{32}P] dCTP	0.05 μl
cDNA sample	0.25 μl
dH_20	8 μl
Total volume	10 μl

The PCR was carried with the following programs (Chapter 2.3.6.1).

2.3.5.2 EXL DNA Polymerase PCR.

EXL DNA Polymerase was used in PCR for APOB exon 26 given the long length of the product (approximately 7.5 Kb product).

Materials:

- 10nM mix dATP, dTTP, dCTP, dGTP (Promega, Southampton, UK)
- EXL® DNA polymerase (5 U/ µl) (Stratagene, CA,USA)
- EXL reaction buffer x10 (Stratagene, CA,USA)
- Dimethyl Sulfoxide (DMSO) (Stratagene, CA,USA)
- Stabilizing Solution (Stratagene, CA,USA)
- Forward + reverse primer (concentration 50µM), (Sigma, Dorset, UK)
- [α-^{32}P] dCTP (Perkin-Elmer, Wellesley, USA)
- dH_20

Method:

The volume of [α-^{32}P] dCTP added per reaction was adjusted according to α-^{32}P activity. The typical reaction mixture was as follows:

EXL reaction buffer (x10)	2.5 µl
dNTPs (10nM conc. dATP, dTTP, dCTP, dGTP)	1.05 µl
Forward primer	0.25 µl
Reverse primer	0.25 µl
EXL® DNA polymerase	0.5 µl
DMSO	0.75 µl
Stabilizing Solution	0.5 µl
0.05 µl (0.5µ Ci) [α-^{32}P] dCTP	0.2 µl
cDNA sample	0.25 µl
dH_20	18.75 µl

Total volume	25 µl

The PCR was carried with the following programs (Chapter 2.3.6.2).

2.3.6 PCR thermocycling conditions.

The number of cycles was adjusted for each oligonucleotide pair to allow amplification in the exponential phase (Chapter 6.2 - appendix).

2.3.6.1 FastTaq DNA polymerase PCR (GAPDH/IR).

		Temp		time/min:sec	
1.	Denaturation	95°c	for	2:00	
2.	Denaturation	95°c	for	0:30	
3.	Annealing	58°c	for	0:30	} repeat 22 times
4.	Extension	72°c	for	0:30	
5.	Remove GAPDH reactions				
6.	Denaturation	95°c	for	0:10	
7.	Annealing	58°c	for	0:30	} repeat 10 times
8.	Extension	72°c	for	0:30	
9.	End				
10.	Remove IR reactions				

2.3.6.2 ExI DNA polymerase PCR program (APOB exon 26).

		Temp		Time/min:sec	
1.	Denaturation	92°c	for	2:00	
2.	Denaturation	92°c	for	0:10	
3.	Annealing	58°c	for	0:30	} repeat 10 times
4.	Extension	68°c	for	7:30	
5.	Denaturation	92°c	for	0:10	
6.	Annealing	58°c	for	0:30	} repeat 20 times
7.	Extension	68°c	for	7:40	
8.	End				

2.3.7 Optimising PCR thermocycling conditions

PCR reactions were run with each primer pair over a range of cycle lengths. The products were then run on a polyacrylamide gel and autoradiographed to determine the optimum number of cycles required to allow amplification in the exponential phase. The subsequent number of cycles per PCR reaction for a given set of primers was fixed accordingly.

2.4 Gel electrophoresis.

2.4.1 Polyacrylamide gel.

Materials:

- Acrylamide solution 30%, (Biosynthesis, Hercules, CA, USA)
- Tetramethylethylenediamine (TEMED) (BDH labs, Poole, UK)
- 10% ammounium persulphate (APS) solution (Sigma, Dorset, UK)
- TBE x1
- dH_20
- Plain and notched glass VGT2 plates
- VGT2 tank

Method:

6-8% polyacrylamide gels were generally run. Gels were made according to the following:

6% Polyacrylamide gel

- 10.5ml dH_20
- 1.5ml TBE x1

- 2.99ml of 30% acrylamide solution
- 5.25 µl TEMED
- 240 µl 10% APS

8% Polyacrylamide gel

- 9.5ml dH_20
- 1.5ml TBE x1
- 3.99ml of 30% acrylamide solution
- 5.25 µl TEMED
- 240 µl 10% APS

Gels were prepared as above with APS solution being added last. Gel solutions were poured between two VGT2 plates and 0.38-0.400mm combs inserted. Gels were allowed to set for 45 minutes. The combs were then removed and the gel was then placed in a VGT2 tank filled with TBE x1. Gels were pre-run for 10 minutes. The products and DNA marker ladders were loaded. Eight percent polyacrylamide gels were run for experiments involving the IR exon 11 at 12W. Six percent gels were run for experiments involving APOB exon 26 at 35W. Once run, gels were dried and autoradiographed and/or exposed to the phosphor-imager plate.

2.4.2 Agarose gel.

Materials:

- Agarose (electrophoresis grade), (Gibco/Invitrogen, Paisley, UK)
- Ethidium Bromide (500µg/ml), (Sigma, Dorset, UK)
- TAE × 50, (National Diagnostics, Atlanta, GA, USA)
- DirectLoad™ Wide Range DNA Marker (Sigma, Dorset, UK)

- Loading dye (Promega, Southampton, UK)

Method:

0.8-1.0% agarose gels were run. A 1.0% agarose gel consisted of 2g agarose, 4ml TAE x50, 10µl ethidium bromide, dH_20 up to 200ml. The gel was melted until liquid and then poured into a cast and a coomb was inserted to construct wells. Once set, the cast was placed in a tank filled with TAE solution (1 litre - 20ml TAE x50, 50 µl ethidium bromide, 980ml dH_20) and the coomb was removed. 4 µl of a sample was generally run with 1 µl of a loading dye. 5 µl of a DirectLoad™ Wide Range DNA Marker ladder was run with samples to determine fragment size. The voltage used across the tank was 150 mV.

2.4.3 Gel imaging.

2.4.3.1 Agarose gels.

Agarose gels were visualized under ultraviolet light (Unvidoc UV Transilluminator) and images were recorded with the Kodak Electrophoresis and Documentation analysis system (Kodak, Herefordshire, UK).

2.4.3.2 Polyacrylamide gels.

2.4.3.2.1 Phospho-imager.

Quantitation of polyacrylamide gels was carried out using a Molecular Dynamics Typhoon PhosphorImager and ImageQuant software (GE Healthcare Life Sciences, Piscataway, NJ). Polyacrylamide gels were dried and exposed to a blanked phosphor-imager plate in a black-out cassette for 1-2 hours depending of the level of exposure required. The ImageQuant software allowed bands to be

quantified and corrected for background counts. The quantified bands were normalized according to the GC content.

2.4.3.2.2 Autoradiography.

Autoradiographs of wet or dry polyacrylamide gels were taken. If wet the gel was first covered with saran wrap prior to being exposed to a radiograph (Kodak Xomat L5 Film, Hertfordshire, UK) placed in a black-out cassette. The cassette was placed at -80^{0}C for between 1-24 hours depending on the level of exposure required. The autoradiograph was developed using a laboratory developer (CompactX4, Xograph Imaging System, UK).

2.5 Immunofluoresence.

Materials:

- 6 chamber slide wells
- PBS x1 (Sigma, Dorset, UK)
- 4% paraldehyde/PBS (Sigma, Dorset, UK)
- 4',6-diamidino-2-phenylindole satin (DAPI), (Sigma, Dorset, UK)
- Fluorescent Mounting Medium (DAKO) (DakoCytomation, Glostrup, Denmark)
- Fluorescent confocal microscope (Zeiss LSM510)

Method:

Cell suspensions were transfected according to Chapter 2.2.7 with fluorescein tagged ASOs and the relevant controls. 300 µl of transfected cell suspension was added to a 6 chamber slide well. These were incubated at 37ºC for 48 hours. Media was then removed and 2 washes with 350 µl of PBS at 4ºC were carried out. Fixation occurred using 350 µl of 4% paraldehyde/PBS placed in each well for ten

minutes. Two further 350 µl PBS washes were carried out. A 2 µg/ml DAPI solution was prepared from stock (10mg/ml). 350 µl DAPI was applied for 30 seconds followed by two 350 µl PBS washes. The plastic chamber and silicon seals were peeled off to leave the slide and fixed cells. One drop of DAKO mounting media was applied to each well and a cover slip applied. The slide was placed in a fluorescent confocal microscope and appropriate images recorded. Slides were stored in the dark.

2.6 DNA sequencing.

2.6.1 Purifying bands.

Materials:

- TE x1 (Gibco/Invitrogen, Paisley, UK)
- 3M NaOAc pH 5.2
- Carrier tRNA, (Qiagen, Valencia, CA)
- Ethanol (VWR international, Lutterworth, UK)

Method:

PCR products were run on a polyacrylamide gel according to protocol (Chapter 2.4.1). The gel was transferred to blotting paper with fluorescent markers applied for positioning, exposed to an autoradiograph and placed at -80ºC for between 4-24 hours depending on the level of exposure required. The film was developed and placed on the thawed gel according to the position of the markers. The bands of interest were identified, cut from the gel with a scapel, and placed in an eppendorf tube. The gel was mashed up in 25 µl of TE x1 and placed at 42ºC for 30 minutes. The supernatant was collected and the process repeated to give a final volume of 50 µl of supernatant. 5 µl of 3M NaOAc pH 5.2, 0.5 µl of carrier RNA and 100 µl of 100% ethanol were added to precipitate the DNA. The mixture was then placed at -

80ºC for 30 minutes. Following this it was spun at 14,000rpm, 4ºC for 15 minutes. The supernatant was removed and 100 µl of 70% ethanol was added and spun at 14,000rpm/4ºC for a further 15 minutes. The supernatant was removed and the DNA sample left to dry. The sample was eluted in 10 µl of TE x1.

2.6.2 DNA sequencing.

PCR product sequencing was carried out at the Genome centre, William Harvey Research Centre. An ABI 3700 automated DNA/RNA Sequencer (Applied Biosystems) was used employing the ABI Prism Big Dye sequencing Kit. Data Collection software 2.0 was used to collect data and results were analysed with the BioEdit sequence analysis software version 7.

2.7 Miscellaneous – buffer solutions/markers/ladders.

2.7.1 Ethylenediaminetetraacetic acid (EDTA) 0.5M pH 8.0.

Materials:

- EDTA, (Sigma, Dorset, UK)
- NAOH, (Sigma, Dorset, UK)
- dH_20

Method:

In order to produce 1 litre solution of 0.5M pH 8.0 EDTA the constituents were mixed as follows:

- EDTA 186.1 g
- dH_20 500 ml
- NAOH added to a pH 8.0

EDTA was dissolved in 500ml of dH_20. NaOH was then added gradually whilst assessing the pH of the solution until it reached pH8.0. The remaining dH_20 was added to make 1 litre.

2.7.2 Tris/Borate/EDTA (TBE) buffer.

Materials:

- Tris base, (Sigma, Dorset, UK)
- Boric acid, (Sigma, Dorset, UK)
- EDTA 0.5M pH 8.0
- dH_20

Method:

To make a x10 solution of TBE buffer, the above mentioned materials were mixed in the following quantities:

- Tris base 108 g
- Boric acid 55 g
- EDTA 0.5M pH8.0 40 ml
- dH_20 960 ml

Once mixed, samples were filtered prior to storage. The solution was diluted 1 in 10 prior to use.

2.7.3 Tris-acetate (TAE) buffer.

Materials:

- TAE x50, (National Diagnostics, Georgia, USA)
- Ethidium bromide, (Sigma, Dorset, UK)
- dH_2O

Method:

The constituents above were mixed in the following quantities to give TAE buffer x1:

- TAE x50 20 ml
- Ethidium bromide stock 50 µl
- dH_20 980 ml

2.7.4 Sodium Acetate (NaOAc) solution (3.0M).

Materials:

- $NaC_2H_3O_2$, (Sigma, Dorset, UK)
- dH_20

Method:

To make a 3.0M Sodium Acetate solution, 246.09g of $NaC_2H_3O_2$ were taken and 800ml of dH_20 added. Once dissolved further water was added to make 1000ml solution.

2.7.4 Ladder markers pBr 322/MSpI or N32325.

Materials:

- Escherichia coli Pol buffer (New England Biolabs, Ipswich, MA, USA)
- pBr 322/MSpI digest (New England Biolabs, Ipswich, MA, USA)

  or

  N32325 (New England Biolabs, Ipswich, MA, USA)
- Klenow buffer x 10 (New England Biolabs, Ipswich, MA, USA)
- dTTP, dATP, dGTP (Promega, Southampton, UK)

- [α-^{32}P] dCTP (Perkin-Elmer, Wellesley, USA)
- dH_20

Method:

5 µl [α-^{32}P] dCTP, 1 µl E.Coli Pol buffer, 1 µl pBr 322/MSpI digest (or N32325) and 1 µl of Klenow buffer (x10) were added to a eppendorf tube and left at room temperature for ten minutes. The 1 µl each of dTTP, dATP and dGTP was added and left at room temperature for a further ten minutes. Then 190 µl of dH_20 was added to make the stock solution. According to the [α-^{32}P] dCTP activity, this stock was diluted with dye and TE to make ladder markers for the polyacrylamide gel.

2.7.5 Bromophenol/xylene cyanolone dye markers.

Materials:

- Sucrose 50%
- EDTA 0.5M, (Sigma, Dorset, UK)
- Tris pH 8.0, (Sigma, Dorset, UK)
- Bromophenol blue, (Sigma, Dorset, UK)
- Xylene cyanolone, (Sigma, Dorset, UK)

Method:

The above constitutes were mixed to give 10mls of dye:

- Sucrose 50% 5 g
- EDTA 0.5M 60mmol 1.2 ml
- Tris pH 8.0 10mmol 100 µl
- Bromophenol blue 0.03% 2.3 mg
- Xylene cyanole 0.03% 2.3 mg

On account of the significant sucrose content, the solution was passed through a bacterial filter to avoid bacterial contamination prior to storage.

2.8 Computational identification of exonic splicing enhancers.

ESEfinder (http://rulai.cshl.edu/tools/ESE/), RESCUE-ESE (http://genes.mit.edu/burgelab/rescue-ese/index.html) and PESXs server (http://cubweb.biology.columbia.edu/pesx/) were used to search the IR exon 11 sequence for exonic splicing enhancers. The FAS-ESS web server (http://genes.mit.edu/fas-ess/) provides predictions for ESS sequences.

ESEfinder 3.0 is a web-based tool utilising the SELEX approach (systematic evolution of ligands by exponential enrichment) to identify ESE motifs for splicing factors; SF2/ASF; SRp55; SRp40 and SC35. SELEX employs an in iterative enrichment process to identify sequences with enhancer properties using an in vitro or in vivo minigene model and provides weight score matrices to assess a given sequence.

RESCUE-ESE uses a hybrid computational/experimental approach to identify hexamers with enhancer characteristics (Fairbrother et al. 2002). Computational analysis identified sequences occurring at high frequency in positions where enhancer sequences may be expected to reside i.e. in exonic rather than intronic sequences and in exons with weaker, non-consensus splice sites rather than exons with strong consensus splice sites which are constitutively spliced. These identified sequences were then experimental tested in minigene reporter constructs to assess their enhancer properties.

PESX Server (Putative Exonic Splicing Enhancers/Silencers) uses a purely bioinformatics process to identify octamer enhancer sequences (Zhang et al. 2004). Octamers from internal noncoding exons (high content of ESE sequences) were compared with sequences predicted to have a low content of ESE sequences i.e.

unspliced pseudo exons and 5′ untranslated regions of transcripts of intronless genes. This method circumvents the 'noise' produced by coding exonic sequences by using non-coding exon sequences.

The FAS-ESS server provides predictions for potential ESS sequences (Wang et al. 2004). The program is based on experiments identifying sequences with silencer activity using an in vivo splicing reporter system based on Green Fluorescent Protein (GFP).

2.9 Statistical analysis.

Statistical analysis and graphing was carried out using Prism 4.0c software (GraphPad Software, San Diego, CA). Three replicates for each experiment were carried out unless stated otherwise. Relevant statistical test were carried out as stated in each experiment (see Chapter 3.0 - Results).

Chapter 3

Results

Chapter 3.1

Modification of insulin receptor exon 11 splicing using antisense oligonucleotides.

3.1.1 A pilot study: 2′-O-methyl RNA ASOs targeting the insulin receptor exon 11.

Background

Exon 11 of the IR pre-mRNA can be alternatively spliced to give two isoforms which vary by exon 11 skipping (isoform IR-A) and inclusion (isoform IR-B). The proportions of these isoforms vary between organs, metabolic states and species. Previous studies have altered exon 11 splicing using deletion studies in minigenes in cell culture (Kosaki et al. 1998) and via hormonally-mediated methods (Kosaki et al. 1993). However these approaches suffer from a lack of physiological relevance or are confounded by various hormonal treatments. ASOs bind to the native pre-mRNA and their mechanism of action is thought to be either via blocking access of trans-acting splicing factors or altering pre-mRNA secondary structure and splicing efficiency. This can cause exon skipping, or the use of an alternative or cryptic splice site.

The aim of the present study was to alter exon 11 splicing using ASOs designed to bind the pre-mRNA by Watson-Crick base-pairing. The prediction was that certain 2′-O-methyl RNA ASOs would enhance either exon skipping or inclusion by targeting key intronic and exonic sites influencing IR exon 11 splicing as guided by the work of Kosaki et al. (Kosaki et al. 1998).

Aim

To change splicing of the IR exon 11 in cell culture using 2′-O-methyl RNA ASOs targeted to key intronic and exonic sites around exon 11 of the IR pre-mRNA.

Methods

The initial ASO chemistry used was 2′-O-methyl RNA. These oligonucleotides contained a phosphate-ribose backbone and a methyl group in place of a hydroxyl group in the 2 position of the ribose ring. This chemistry was chosen because of the increased binding capacity to pre-mRNA and resistance to RNase mediated degradation. 2′-O-methyl ASOs were ordered from Dharmacon (Lafayette, CO, USA). They were designed to bind and occupy pre-mRNA sequences by Watson-Crick base-pairing to the INSR sequence as taken from the Ensembl Genome Browser (http://www.ensembl.org/index.html). ASOs were either 18 or 36 nucleotides in length, in-line with previous studies, such that targeted sequences were adequately covered (Khoo et al. 2007). ASO sequences are given in Chapter 6.1 - appendix. These experiments were carried out at a point when the key splicing factors binding sites within exon 11 had not yet been fully elucidated (see Chapter 4.2). 2′-O-methyl RNA ASOs were designed complementary to single intronic sites, 18 nucleotides long (figure 3.1.1.A); the 3′ splice site (designated INSR 3); the 5′ splice site (designated INSR 5) and the branch point sequence (designated INSR B). The ASOs binding the 3′ and 5′ splice sites were designed without overlapping the intron/exon boundaries. ASOs which targeted two intronic sites within the same oligonucleotide, 36 nucleotides long, were also synthesised. These included ASOs to the 3′ splice site and branch point sequence (designated INSR 3B); the 5′ splice site and the branch point sequence (designated INSR 5B) and the 3′ splice site and 5′ splice site (designated INSR 53).

Exonic-targeted ASOs, 18 nucleotides long, were designed (figure 3.1.1.B) to bind the 5′ half of exon 11 (designated INSR ex1) and the 3′ half of exon 11 (designated INSR ex2). An ASO, 36 nucleotides long, was designed to block the whole of exon 11 (designated INSR ex1+2).

Following initial experiments with the above-mentioned sequences, two further 2′-O-methyl RNA ASOs were designed and subsequently tested (figure 3.1.1.C). These included the ASO INSR 6/7, which was a 43 nucleotide sequence targeting an intron 10 sequence, 52 nucleotides upstream of the 3′ splice site. This sequence has been identified as having an inhibitory influence on IR exon 11 splicing from a previous study in which it has been identified as sequence 6/7 (Kosaki et al. 1998). A further exonic sequence was targeted by the ASO INSR ex3 which was an 18mer targeting the central portion of IR exon 11. INSR ex3 was tested to further delineate influential sites in the IR exon 11.

Detailed methodology is provided in Chapter 2 - materials and methods. Essentially, ASOs were reverse-transfected into cultured HepG2 cells using the transfection agent Lipofectamine 2000 (Gibco/Invitrogen, Paisley, UK). ASOs were transfected at a concentration of 250 nM, in keeping with similar previous studies (Khoo et al. 2007). Dose titration experiments were also carried out to establish a dose-response relationship (Chapter 3.1.3). ASOs were incubated with HepG2 cells for 48 hours at 37^{o}C/5% CO_2. HepG2 cells were used as they represent a well-established model of IR expression and action. They also retain much of normal hepatocyte function (Javitt 1990). RNA was subsequently harvested and subjected to RT-PCR. The PCR primers allowed identification of the 2 isoforms of the IR and GAPDH run as a control for mRNA quantity and quality. Products were run on an 8% polyacrylamide gel. The PCR incorporated the radioactive nucleotide [α-^{32}P] dCTP (Perkin-Elmer, Wellesley, USA) allowing relevant bands to be

identified and quantified using a Phosphorimager. Radioactive counts were corrected for GC content. The effects of the ASOs were expressed by measuring the amount of IR-B (IR isoform including exon 11) as a percentage of total IR (IR-A + IR-B).

The GAPDH control was included to demonstrate appropriate loading of samples and a marker of RNA quality but this was not used to correct for well loading. The reason for this was that the figure being compared (IR-B/ (IR-A+IR-B)) for each column was a proportional rather than an absolute figure which would eliminate differences according to sample loading if corrected for GAPDH. If the sample loaded in a particular column happened to be lower then IR-B and IR-A+IR-B would be proportionally reduced by the same amount thus eliminating any effect. This assumption was corroborated by calculations from initial experiments which demonstrate the percentage IR-B with respect to total IR was identical whether or not it was corrected for GAPDH.

Statistical analysis was carried out using a one-way repeat ANOVA to assess for overall statistical significance between the groups. A Dunnett's multiple comparison test was used to assess for statistical differences between the proportion of exon 11 inclusion in the no-oligo control cells compared with cells transfected with each of the ASOs (significance $p<0.05$). Each experiment was repeated in triplicate.

An analysis of the IR exon 11 using web-based tools (Chapter 2.8) capable of identifying sequences with ESS and ESE activity was undertaken to compare the findings generated from this study.

Figure 3.1.1.I shows data from equivalence studies for RNA quantification comparing samples using the NanoDrop Spectrophotometer with the Ribogreen QuantIT estimation.

Figure 3.1.1.A

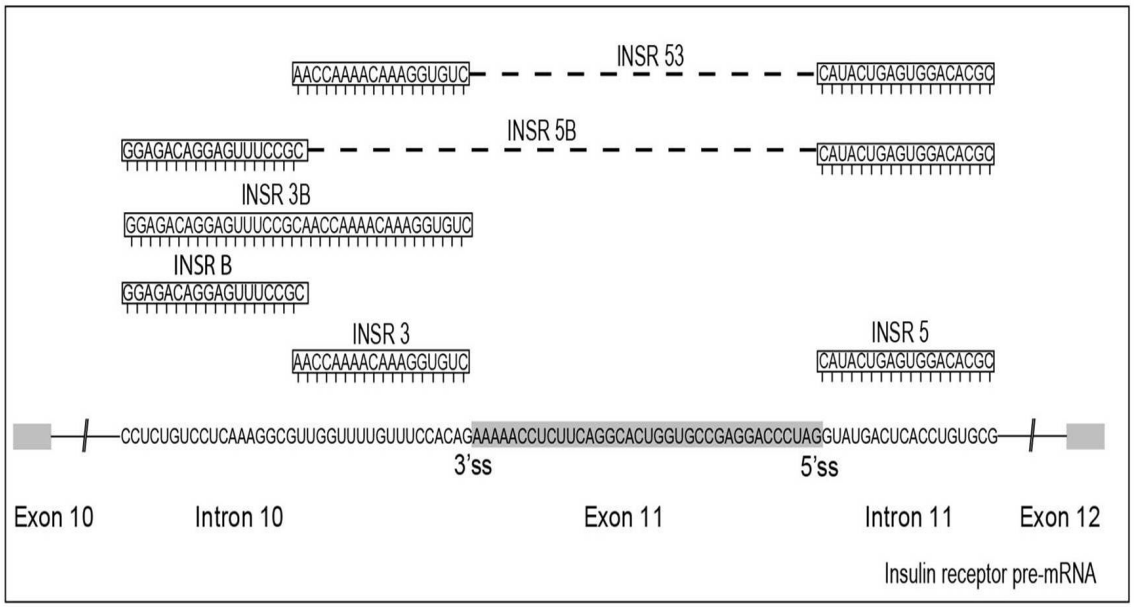


Figure 3.1.1.A Intronic ASOs targeting targeting regions surrounding the IR. Schematic diagram shows a portion of the IR pre-mRNA from exon 10 to exon 12 (not to scale). Exon 11 and surrounding intronic sequences are shown. Exons are delineated by the grey shading, the rest of the sequence being intronic. ASOs are illustrated above the IR pre-mRNA. ASOs, 18 nucleotides long, were designed to block the 3′ splice site (designated INSR 3), the 5′ splice site (INSR 5), the branch point sequence (INSR B). Combinations of these ASOs, 36 nucleotides long, were also designed to block two sites simultaneously. These included the 3′ splice site and the branch point sequence (designated INSR 3B), the 5′ splice site and the branch point sequence (INSR 5B) and the 3′ splice site and the 5′ splice site (INSR 53). The dotted line schematically represents sequences in apposition to each other rather than a linker sequence between them. Cells were transfected with ASOs at varying concentrations. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.

Figure 3.1.1.B

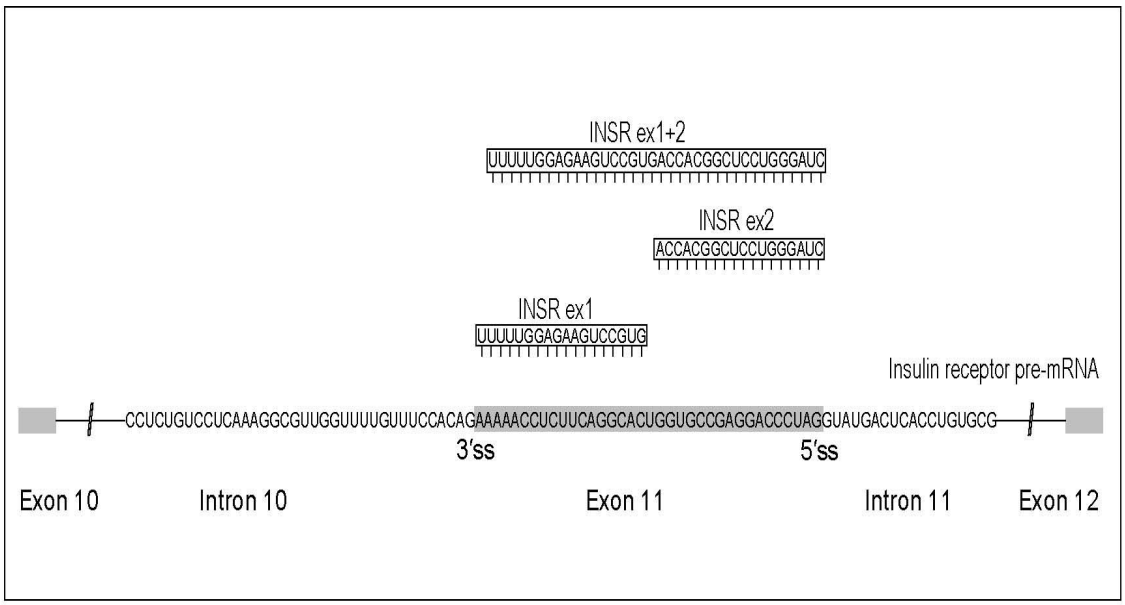


Figure 3.1.1.B Exonic ASOs targeting the IR exon 11. Schematic diagram shows a portion of the IR pre-mRNA from exon 10 to exon 12 (not to scale). IR exon 11 and surrounding intronic sequences are shown. Exons are delineated by the grey shading, the rest of the sequence being intronic. ASOs are illustrated above the IR pre-mRNA. ASOs, 18 nucleotides long, were designed to block the two halves of exon 11 i.e. INSR ex1 and INSR ex2. A combination of these ASOs, designated INSR ex1+2 (36 nucleotides long) blocked both sites simultaneously. Cells were transfected with ASOs at varying concentrations. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.

Figure 3.1.1.C

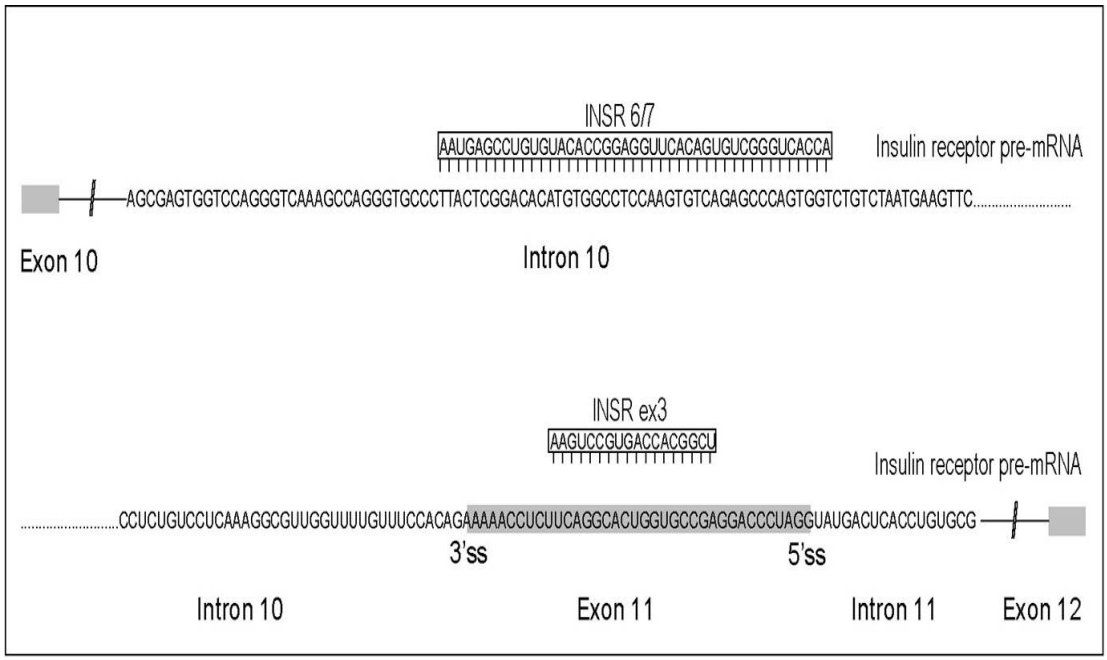


Figure 3.1.1.C 2′-O-methyl RNA ASOs targeting intron 10 and exon 11 of the IR pre-mRNA. Schematic diagram shows a portion of the IR pre-mRNA from exon 10 to exon 12 (not to scale). Exon 11 and surrounding intronic sequences are shown. Exons are delineated by the grey shading, the rest of the sequence being intronic. ASOs are illustrated above the IR pre-mRNA. An ASOs 18 nucleotides long, designated INSR ex3, was designed to block the central portion of the IR exon 11. An intronic ASO targeting a 43 nucleotide sequence in intron 10, 52 nucleotides upstream of the 3′ splice site of exon 11 was designed. Cells were transfected with ASOs at varying concentrations. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.

Figure 3.1.1.D

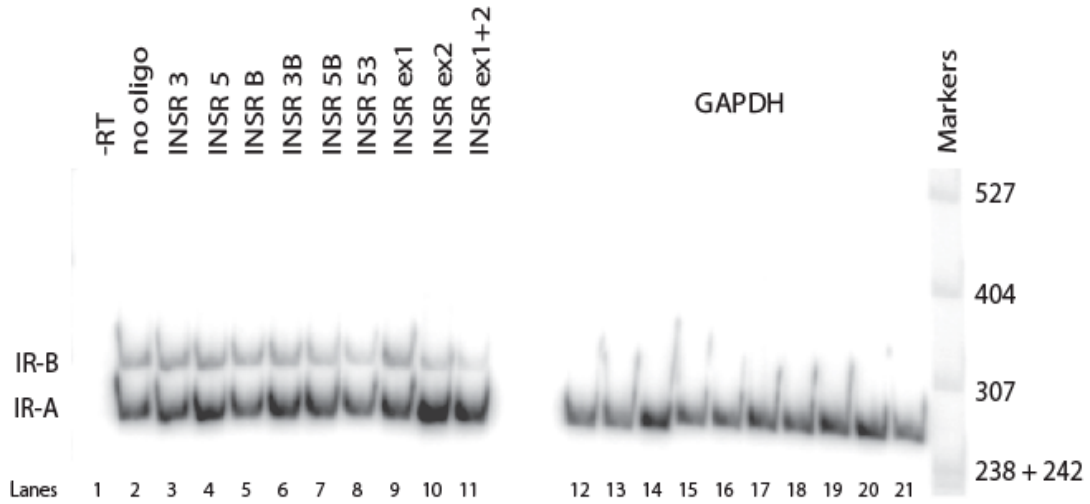


Figure 3.1.1.D IR exon 11 splicing in the presence of targeted 2′-O-methyl RNA ASOs. Polyacrylamide gel from a representative experiment showing IR isoforms from HepG2 cells transfected with ASOs. Each column represents a well of transfection with ASOs marked above the column. The left side of the figure represents the two IR isoforms - IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion). The right side of the figure represents GAPDH run as a control for mRNA quantity and quality.

Figure 3.1.1.E

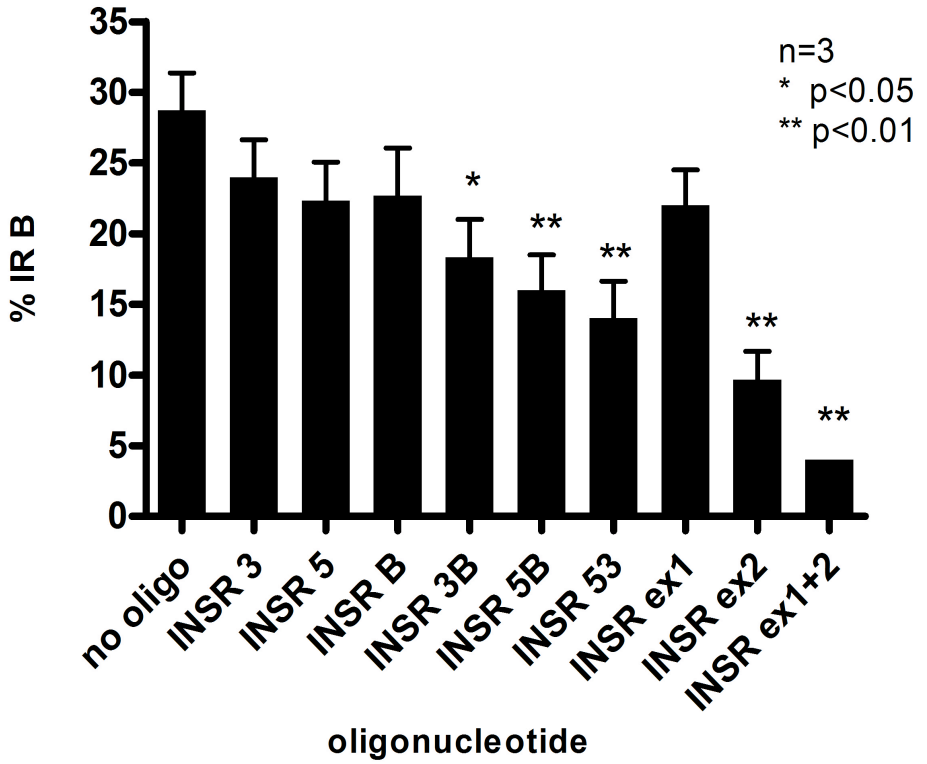


Figure 3.1.1.E IR exon 11 splicing in the presence of targeted 2′-O-methyl ASOs. The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection by ASOs in HepG2 cells at a concentration of 250nM. Quantitative data from three independent replications of the experiment (error bars show the S.E.M.). Overall one-way repeated measures ANOVA p-value was <0.0001, indicating statistically significant differences between groups. Dunnett's multiple comparison test was used to compare the no-oligo control versus ASO transfected cells: statistically significant differences were found with ASO INSR 3B, INSR 5B, INSR 53, INSR ex2 and INSR ex1+2. * indicates $p<0.05$; ** indicates $p<0.01$.

Figure 3.1.1.F

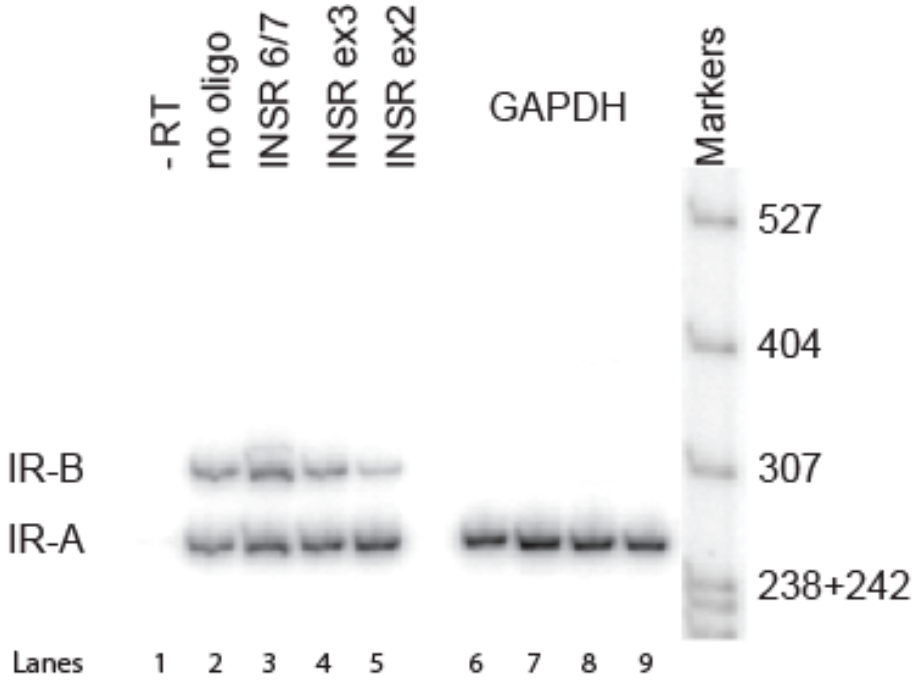


Figure 3.1.1.F IR exon 11 splicing in the presence of the 2′-O-methyl RNA ASOs INSR 6/7 and INSR ex3. Polyacrylamide gel showing IR isoforms from cells transfected with 2′-O-methyl RNA ASOs INSR 6/7 and INSR ex3 - IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion). 2′-O-methyl RNA ASO INSR ex2 was run as a positive control. GAPDH was run as a control for mRNA quantity and quality.

Figure 3.1.1.G

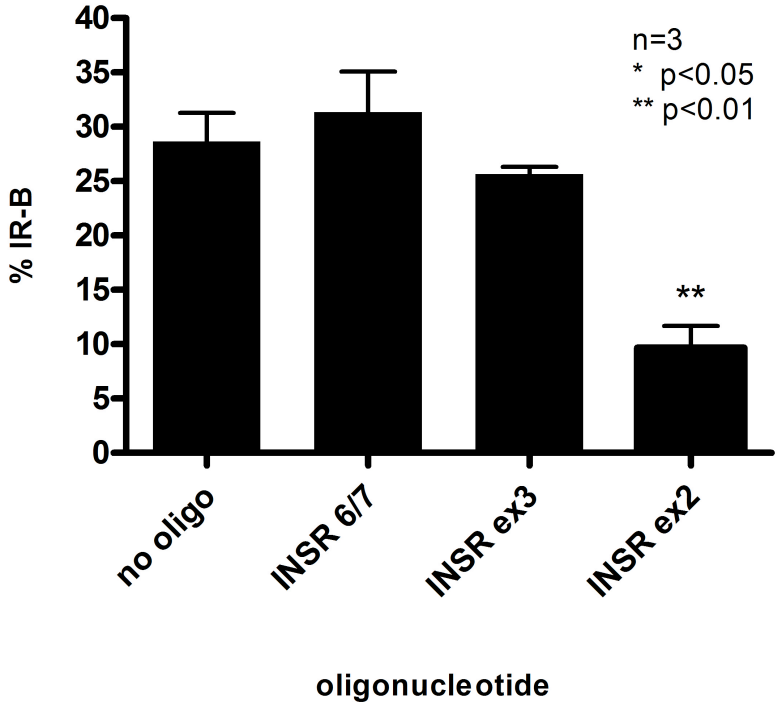


Figure 3.1.1.G Alternative splicing of the IR exon 11 induced by 2′-O-methyl RNA ASOs INSR 6/7 and INSR ex3. The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection by 2′-O-methyl RNA ASOs INSR 6/7 and INSR ex3 at a concentration of 250nM. Quantitative data from three independent replications of the experiment (error bars show the S.E.M.). Overall one-way repeated measures ANOVA p-value was <0.0001 indicating statistically significant differences between groups. Dunnett's multiple comparison test was used to compare the no-oligo control versus 2′-O-methyl RNA ASOs transfected cells. Neither INSR 6/7 or INSR ex3 were effective at altering IR splicing. The 2′-O-methyl RNA ASO INSR ex2, run as a positive control, was the only effective ASO.

Table 3.1.1.A

Program	ESS/ESE	Position in IR exon 11	Sequence
ESE Finder	SRp40 (ESE)	9-15	CTTCAGG
		15-21	GCACTGG
	SC35 (ESE)	28-35	GGACCCTA
		29-36	GACCCTAG
	SF2/ASF(ESE)	16-22	CACTGGT
		21-27	GTGCCGA
		24-30	CCGAGGA
PESX	ESE	8-15	TCTTCAGG
		25-34	CGAGGACCCT
	ESS	-	-
RESCUE-ESE	ESE	2-7	AAAACC
		3-8	AAACCT
		8-13	TCTTCA
		9-14	CTTCAG
FAS-ESS	ESS	-	-

Table 3.1.1.A Predicted enhancer/silencer sites in the IR exon 11. Column 1 lists the web-based tools used to identify potential enhancer and silencer elements in the IR exon 11. URLs for these tools are provided in Chapter 6.3 - appendix. Column 2 gives the type of element (ESE or ESS) identified. ESE finder specifies several ESEs including SRp40, SC35 and SF2/ASF. Column 3 provides the location for the given ESS or ESE with exon 11, numbering between 1-36 nucleotides from the 5′ end of the exon. Column 4 gives the precise sequences of the predicted ESSs and ESEs.

Figure 3.1.1.H

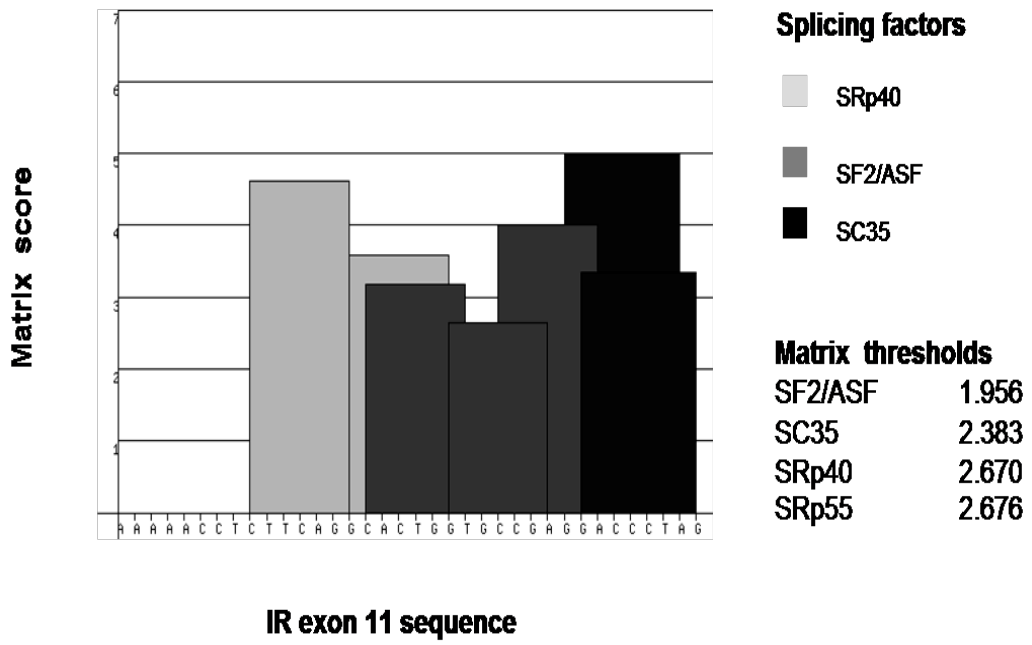


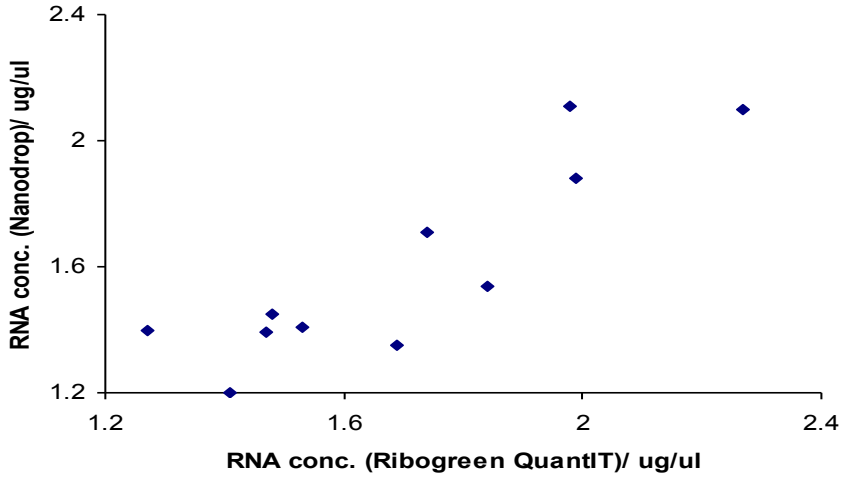
Figure 3.1.1.H Predicted ESEs in the IR exon 11- generated by the ESE finder analysis. (Chapter 6.3 - appendix for URL) Graphical representation of predicted ESE sites in the IR exon 11 for SF2/ASF, SC35, SRp40 and SRp55. IR exon 11 sequence given on the x axis, matrix score (predicted strength of ESE) given on the y axis. ESE are predicted in the centre and to the 3′ half of IR exon 11. The strongest prediction is that for a SC35 binding site at position 24-30 (matrix score 4.99).

Figure 3.1.1.I

A

Sample	RNA concentration µg/µl		Sample	RNA concentration µg/µl	
	Ribogreen QuantIT	Nanodrop		Ribogreen QuantIT	Nanodrop
1	1.98	2.11	7	1.74	1.71
2	1.53	1.41	8	1.47	1.39
3	1.69	1.35	9	1.48	1.45
4	1.99	1.88	10	1.27	1.40
5	1.84	1.54	11	1.41	1.20
6	2.27	2.10			

B



C

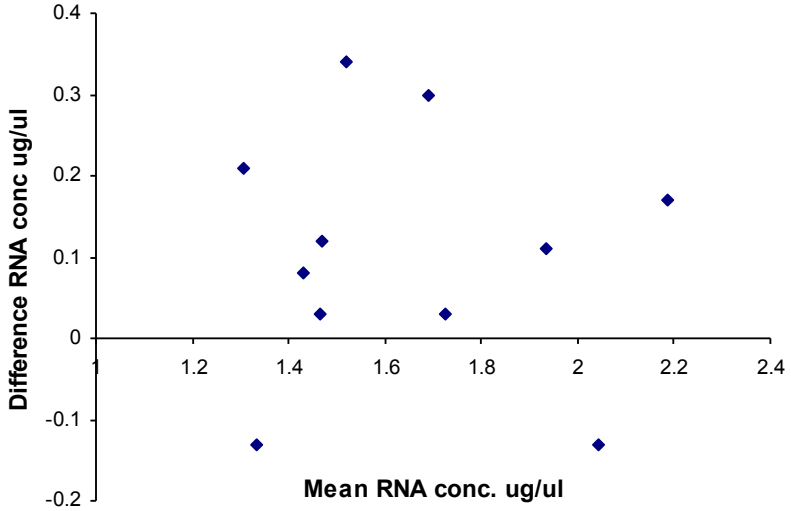


Figure 3.1.1.I RNA quantification equivalence studies - NanoDrop ND-1000 UV-Vis Spectrophotometer vs. Ribogreen QuantIT estimation. A. RNA concentrations measured with two methods. **B.** Scatterplot of RNA concentrations measured by two methods. The calculated correlation co-efficient was 0.88. **C.** Bland-Altman plot of mean RNA concentration against difference in RNA concentration (bias) for each pair of readings (Ribogreen QuantIT – Nanodrop). Mean bias was 0.10 ug/ul (towards Ribogreen QuantIT) with a 95% limit of agreement -0.20 - 0.40.

Discussion

Figure 3.1.1.D and 3.1.1.F show representative gels from experiments involving the transfection of 2′-O-methyl RNA ASOs into cultured HepG2 cells. In native HepG2 cells without the addition of ASOs (figure 3.1.1.D and 3.1.1.F - lane 2 designated 'no oligo') there is a predominance of exon 11 skipping (IR-A) over exon 11 inclusion (IR-B). The proportion of IR-B with respect to the total IR (IR-A + IR-B) was approximately 30% in keeping with reported proportions where HepG2 have not been manipulated with the use of minigenes containing IR transcripts (Kosaki et al. 1993).

Data from three experimental replicates was combined and represented graphically in figure 3.1.1.E and 3.1.1.G. Data was represented as the change in the proportion of exon 11 included isoform with respect to the total IR (exon 11 included isoform + exon 11 skipped isoform) for each ASO transfected.

Intron-targeted ASOs against single sites (INSR 3, INSR 5, INSR B) did not significantly alter exon 11 splicing (figure 3.1.1.D - lanes 3, 4, 5 respectively). This includes the long ASO INSR 6/7 (figure 3.1.1.F - lane 3) targeted to a 43 nucleotide intronic site. Combination ASOs, blocking two intronic sites simultaneously (INSR 3B, INSR 5B and INSR 53) and at the same concentration as ASOs targeting single sites (figure 3.1.1.D - lanes 6, 7 and 8 respectively), significantly increased exon 11 skipping (i.e. reduce proportion IR-B) in the order INSR 53>INSR 5B>INSR 3B. The intronic combination ASOs which spanned the exon (INSR 5B, INSR 53) were more effective ($p<0.01$) than the ASO INSR 3B ($p<0.05$) which did not.

Exonic ASOs INSR ex2 and INSR ex1+2 (figure 3.1.1.D - lanes 10 and 11) were the most effective at causing exon 11 skipping ($p<0.01$), with INSR ex1+2 being

more potent than INSR ex2. INSR ex1 (figure 3.1.1.D - lane 9) and INSR ex3 (figure 3.1.1.F - lane 4) had no significant effect on splicing.

The lack of efficacy of 2′-O-methyl RNA ASOs targeted to single intronic sites around the IR exon 11 may be ascribed to either technical factors in the methodology or a factor inherent to ASOs targeted to single intronic sites.

Technical factors include issues relating to transfection efficiency and ASO design. Transfection efficiency was not formally assessed in this study. However, ASOs transfected in identical conditions were noted to have significant effects on IR exon 11 splicing e.g. INSR ex2. It is possible that different ASOs had different transfection efficiency rates. Variations in transfection efficiency could explain the differential effects of these ASOs. Against this are the facts that the effective ASO INSR ex2 also had a similar design (2′-O-methyl RNA backbone) to the ineffective ASOs targeted to single intronic sites (INSR 3, INSR 5, INSR B). The 18mer INSR ex2 was more effective that the 36mers INSR 3B, INSR 5B and INSR 53 but less effective than the 36mer, INSR ex1+2. It would therefore be unlikely that ASO length was a predominant factor for poor efficacy of ASOs targeting single intronic sites.

If the lack of efficacy related to a factor inherent to ASOs targeted to single intronic sites it suggested there was functional redundancy in the system i.e. blocking of single intronic sites was not rate-limiting to the splicing of IR exon 11. This is in contrast to the finding that ASOs which targeted two intronic sites simultaneously appeared to enhance IR exon 11 skipping.

Exonic ASOs spanning the whole exon and the 3′ half of the IR exon 11 both appeared to cause maximal skipping of exon 11. Those targeting the 5′ and central portion of the exon (INSR ex1 and INSR ex3) had no overall effect on splicing. This data would therefore suggest the 3′ end of exon 11 harbours an enhancer

sequence important for exon inclusion. ASOs spanning the IR exon 11, whether targeting intronic (INSR 5B and INSR 53) or exonic sites (INSR ex1+2) were more effective at causing exon 11 skipping than 36mer ASOs that did not span the exon (INSR 3B). Both the 5B and 53 ASO appear to cause IR exon 11 skipping. In terms of binding, three possibilities exist; the ASOs are binding to two sites simultaneously (5' splice site and 3' splice site or BPS); the ASOs bind only one of the two sites they target (5' splice site, 3' splice site or BPS), in other words only half the intended target (50%) is blocked, or a combination of these two possibilities. If the ASOs are binding the two sites simultaneously then the assumption is that they are blocking the exon. However the fact that these ASOs are not as effective as INSR ex1+2 suggests that they are not blocking the exon in the same way. On the basis of current data it is not possible to speculate beyond this.

While these experiments demonstrated that IR exon 11 skipping could be induced easily, no ASO tested caused increased exon 11 inclusion. The possible explanations for this finding was that there were no silencer elements in the areas targeted, or that the ASOs were unable to block any silencer elements due to a methodological problem.

Using a range of computational and experimental modelling programmes available via the web, the IR exon 11 was analysed to identify possible enhancer/silencer binding sites (table 3.1.1.A and figure 3.1.1.H). Tools identified predicted ESE sites through-out the IR exon 11. ESE finder identified potential binding sites for the splicing factors SRp40, SF2/ASF and SC35 which were found in the centre and 3′ end of exon 11 (figure 3.1.1.H). RESCUE-ESE identified sites predominantly at the 5′ end and centre of exon 11 while PESX predicted enhancer sequences both in the 5′ half (position 8-16) and the 3′ half of exon 11 (position 25-34). Data from the

current study suggests a possible enhancer region in the 3′ half of exon 11 in keeping with RESCUE-ESE and PESX. Tools capable of predicting ESS sites (PESX and FAS-ESS) did not predict any ESS sites in exon 11 which may explain the inability of ASOs to cause IR exon 11 inclusion.

In summary, the findings of these initial experiments fulfil the aim by demonstrating that 2′-O-methyl RNA ASOs targeting key sites in and around the IR exon 11 can alter its splicing. Importantly however, while various 2′-O-methyl RNA ASOs were able to cause IR exon 11 skipping, no ASO was identified which could enhance exon 11 inclusion.

3.1.2 Are the ASOs specific? A study of 2′-O-methyl RNA ASOs with scramble sequences.

Background

The use of ASO technology for both experimental and therapeutic purposes relies on specific ASO binding. ASOs may cause non-specific effects by inadvertently binding similar sequences in other pre-mRNAs. Non-specific ASO binding may be minimised by designing longer ASOs, thereby reducing the chance of matches elsewhere in the genome. ASO may also be checked against other genomic sequences using web-based bioinformatics tools e.g. 'BLAST' (http://blast.ncbi.nlm.nih.gov/blast/Blast.cgi) to avoid matches.

Are the effects of 2′-O-methyl RNA ASOs targeting the IR exon 11 specific? This question can be addressed in two ways. The effect of ASO scrambles, with the same nucleotide composition but randomly sequenced, on the splicing of the IR exon 11 could be tested. Alternatively the effect of ASOs targeting the IR exon 11 on other alternatively spliced exons from different genes could be assessed.

The aim of the following experiments was to confirm that 2′-O-methyl RNA ASOs effects on splicing of the IR exon 11 were targeted and specific. The prediction was that scramble ASOs based on effective 2′-O-methyl RNA ASOs (chapter 3.1.1) would have no effect on IR exon 11 splicing.

Aim

To assess the effect of 2′-O-methyl RNA ASO scrambles on the IR exon 11 splicing.

Methods

ASOs were designed using a web-based tool (www.sirnawizard.com) to scramble sequences. These included scrambles of the effective targeted ASOs INSR 53, designated INSR 53scram; ASO INSR ex2, designated INSR ex2scram and ASO INSR ex1+2, designated INSR ex1+2scram. A scramble sequence of the non-effective ASO INSR ex1, designated INSR ex1scram was also constructed as a further control. BLAST scans (http://blast.ncbi. nlm.nih.gov/blast/Blast.cgi) were carried out to confirm the absence of sequence homology. Scramble ASOs were transfected into HepG2 cells at a concentration of 250 nM as described in Chapter 2 - materials and methods. Effective ASOs identified in Chapter 3.1.1 were also transfected as positive controls. ASOs were incubated with HepG2 cells for 48 hours at 37^{o}C/5% CO_2. RNA was harvested and subjected to RT-PCR for the IR and GAPDH as a control for mRNA quantity and quality. Products were run on a polyacrylamide gel and radioactivity quantified using a Phosphorimager. Each experiment was replicated in triplicate.

Results

Figure 3.1.2.A

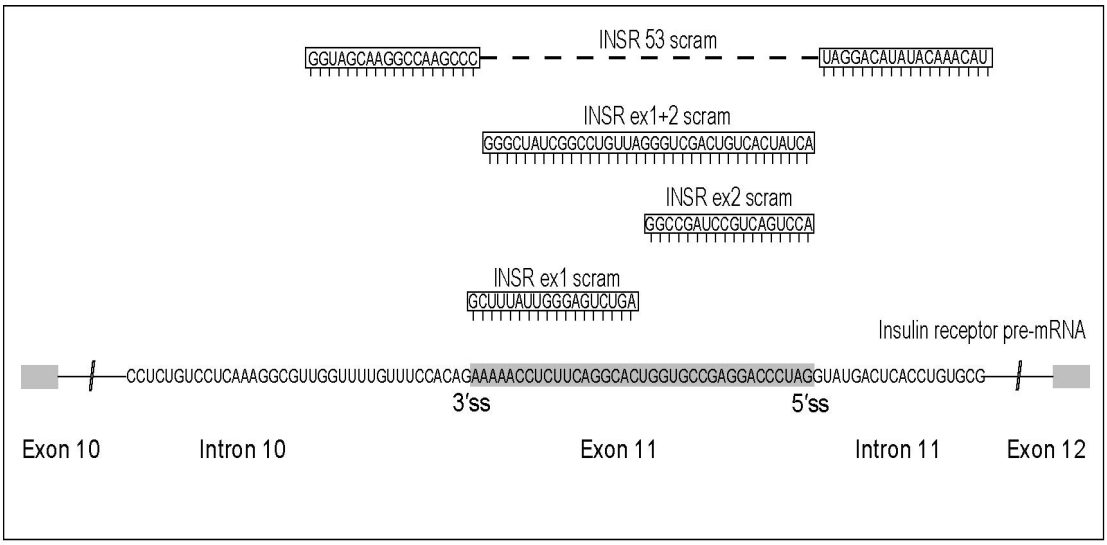


Figure 3.1.2.A Scramble ASOs targeting the IR exon 11. Schematic diagram shows a portion of the IR pre-mRNA from exon 10 to exon 12 (not to scale). Exon 11 and surrounding intronic sequences are shown. Exons are delineated by the grey shading, the rest of the sequence being intronic. ASOs are illustrated above the IR pre-mRNA. Scramble ASOs (INSR 53scram, INSR ex1scram, INSR ex2 scram, INSR ex1+2scram) were designed to assess ASO specificity. The dotted line schematically represents sequences in apposition to each other rather than a linker sequence between them. Cells were transfected with scramble ASOs at 250 nM. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.The scramble ASOs were transfected alongside the non-scramble ASOs; INSR 53; INSR ex1; INSR ex2 and INSR ex1+2 for comparison (not shown in figure).

Figure 3.1.2.B

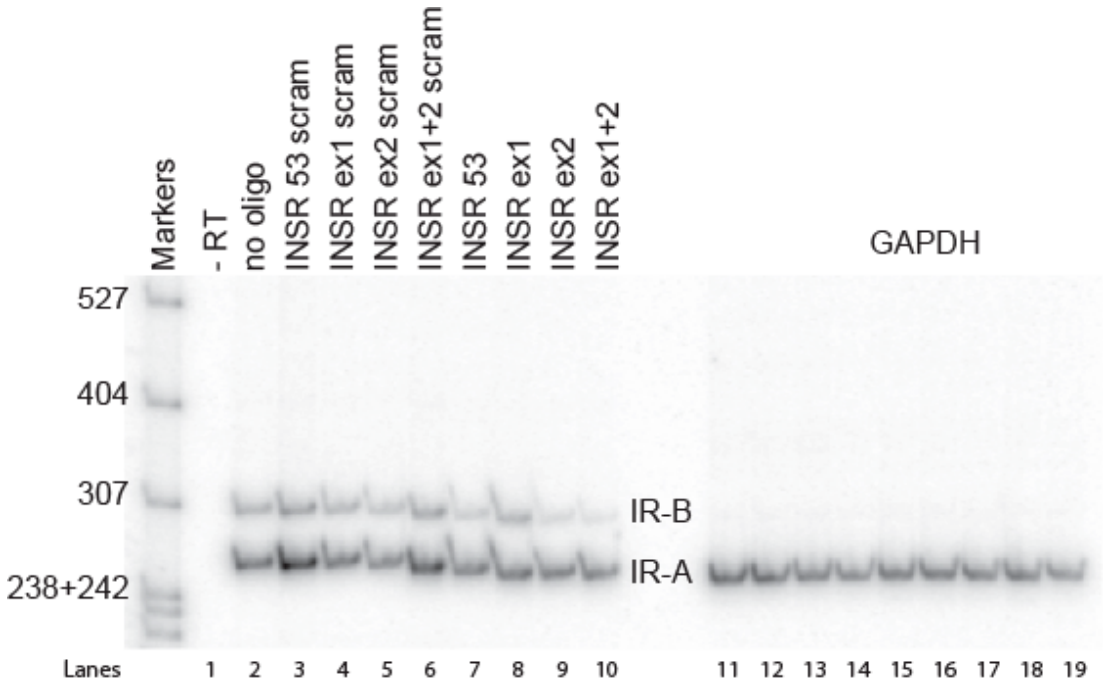


Figure 3.1.2.B IR exon 11 splicing in the presence of 2′-O-methyl RNA ASO scrambles. Polyacrylamide gel from a representative experiment showing IR isoforms from HepG2 cells transfected with ASOs. Each column represents a well of transfection with ASOs marked above the column. The left side of the figure represents the two IR isoforms - IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion). ASO INSR 53, INSR ex1, INSR ex2, INSR ex1+2 were run as positive controls. The right side of the figure represents GAPDH run as a control for mRNA quantity and quality.

Figure 3.1.2.C

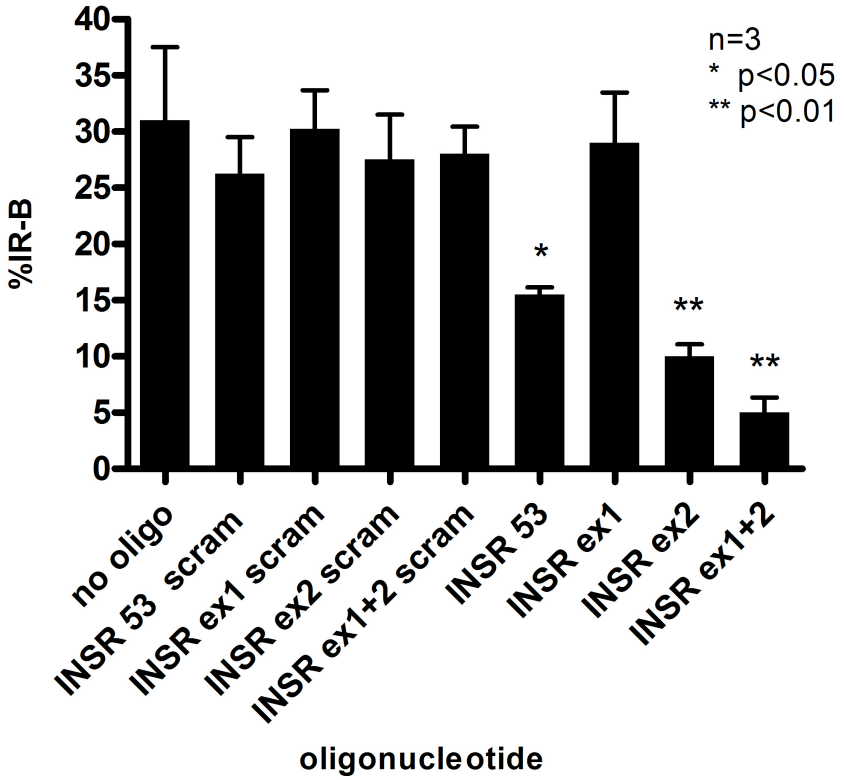


Figure 3.1.2.C IR exon 11 splicing in the presence of 2′-O-methyl RNA ASO scrambles. The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection by 2′-O-methyl ASOs and ASO scrambles at a concentration of 250nM. Quantitative data collated from three independent replications of the experiment (error bars show the S.E.M.). Overall one-way repeated measures ANOVA of the dataset showed $p<0.0001$, indicating statistically significant differences between groups. Dunnett's multiple comparison test was used to compare the no-oligo control versus ASO transfected cells: statistically significant differences were found with ASO INSR 53, INSR ex2 and INSR ex1+2 but no effect on splicing was noted with ASO scrambles.

Discussion

The effect of scramble ASOs on IR exon 11 splicing is demonstrated in a representative gel in figure 3.1.2.B and graphically in figure 3.1.2.C providing the cumulative data from three replicate experiments.

INSR 53scram, INSR ex1scram, INSR ex2scram and INSR ex1+2scram (lanes 3, 4, 5 and 6 respectively) did not alter the proportion of IR-B following transfection. No differences were noted between scramble ASOs of different lengths, from 18mers (INSR ex1scram and INSR ex2scram) to 36mers (INSR 53scram and INSR ex1+2scram). Positive control ASOs INSR 53, INSR ex2 and INSR ex1+2 (lanes 7, 9, 10) all cause exon skipping while INSR ex1 (lane 8) had no overall effect, in keeping with previous data (Chapter 3.1.1). The data shown for ASOs INSR 53, INSR ex2, INSR ex1+2 and ASO INSR ex1 represents 3 separate replicates experiments and is separate from data shown in Figure 3.1.1.E.

These results confirm the prediction that scramble ASOs had no effect on IR exon 11 splicing suggesting that 2′-O-methyl RNA ASOs effects on splicing of the IR exon 11 were specific. Further confirmation of ASO specificity could be obtained by demonstrating the lack of efficacy of 2′-O-methyl RNA ASOs targeting the IR exon 11 on other alternatively spliced genes (see Chapter 4.3 - discussion).

3.1.3. Optimising the concentration of 2′-O-methyl RNA ASOs via titration.

Background

The results from Chapter 3.1.1 and 3.1.2 have confirmed that ASOs targeting the IR exon 11 can influence splicing and that these effects are sequence specific. Are these effects dose-related? Do ASOs identified to be ineffective in Chapter 3.1.1, remain so at varying concentrations? Below optimum concentrations, ASOs would have little effect on splicing and above which there may be cell toxicity. The prediction was that effective ASOs would alter splicing in a dose-related manner while ineffective ASOs would not.

Aim

To confirm the dose-related effects on splicing of ASOs targeting the IR exon 11.

Methods

Detailed methodology is provided in Chapter 2 - materials and methods. The effective ASOs INSR 53 and INSR ex1+2 along with the ineffective ASO INSR ex1 were used in this set of experiments. ASOs were reverse-transfected into HepG2 cells at concentrations between 10-500 nM. Serial dilution using Opti-MEM I (Gibco/Invitrogen, Paisley, UK) were undertaken to achieve this. ASOs were incubated with HepG2 cells for 48 hours. Cells were assessed by microscopy. RNA was harvested and subjected to RT-PCR to identify the two isoforms of the IR and GAPDH as a control for mRNA quality and quantity. Each experiment was repeated in triplicate.

Figure 3.1.3.A

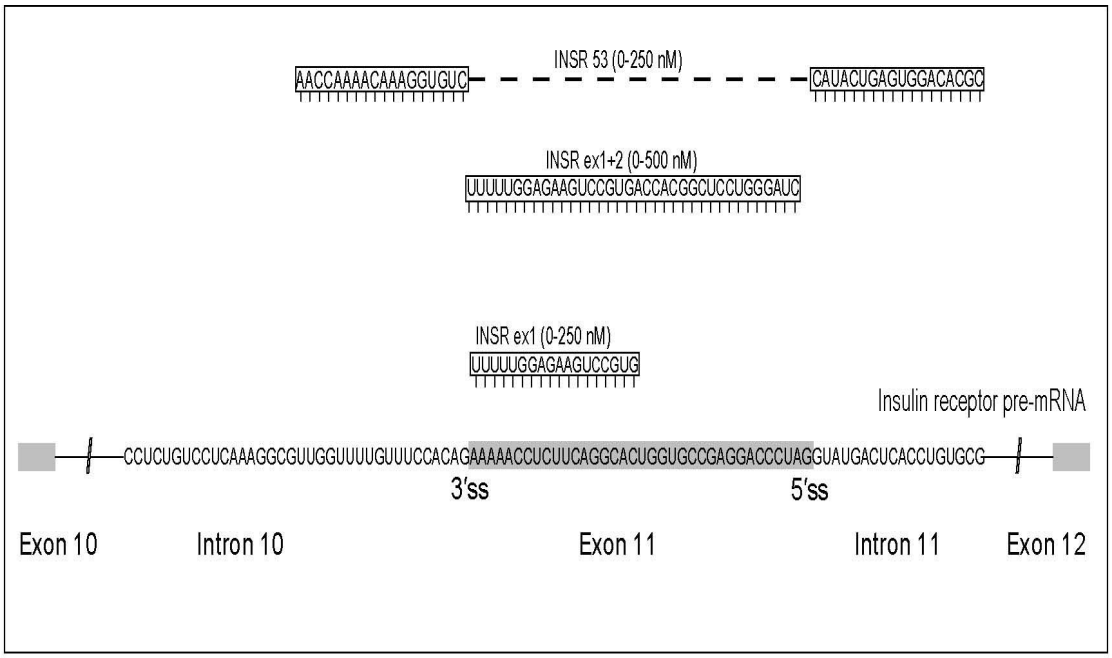


Figure 3.1.3.A Titration of ASOs targeting the IR exon 11. Schematic diagram shows a portion of the IR pre-mRNA from exon 10 to exon 12 (not to scale). Exon 11 and surrounding intronic sequences are shown. Exons are delineated by the grey shading, the rest of the sequence being intronic. ASOs are illustrated above the IR pre-mRNA. INSR ASOs including INSR 53, INSR ex1 and INSR ex1+2 were transfected at increasing concentrations into HepG2 cells. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.

Results

Figure 3.1.3.B

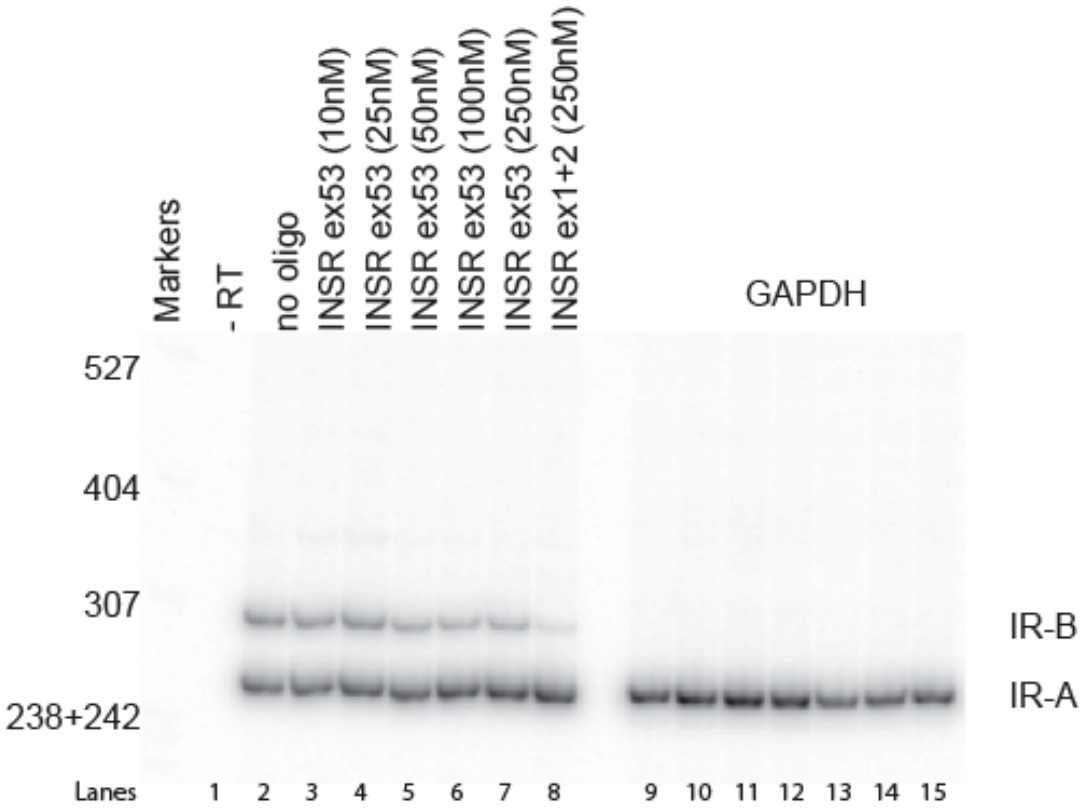


Figure 3.1.3.B Titration of 2′-O-methyl RNA ASO INSR 53 targeting the IR exon 11. Polyacrylamide gel showing IR isoforms from cells transfected with increasing concentrations of ASO INSR 53 (10-250 nM). IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion) marked in the figure relate to the insulin receptor isoforms (left-side), not GAPDH. The ASO INSR ex1+2 was run as a positive control. The right side of the figure represents GAPDH run as a control for mRNA quantity and quality.

Figure 3.1.3.C

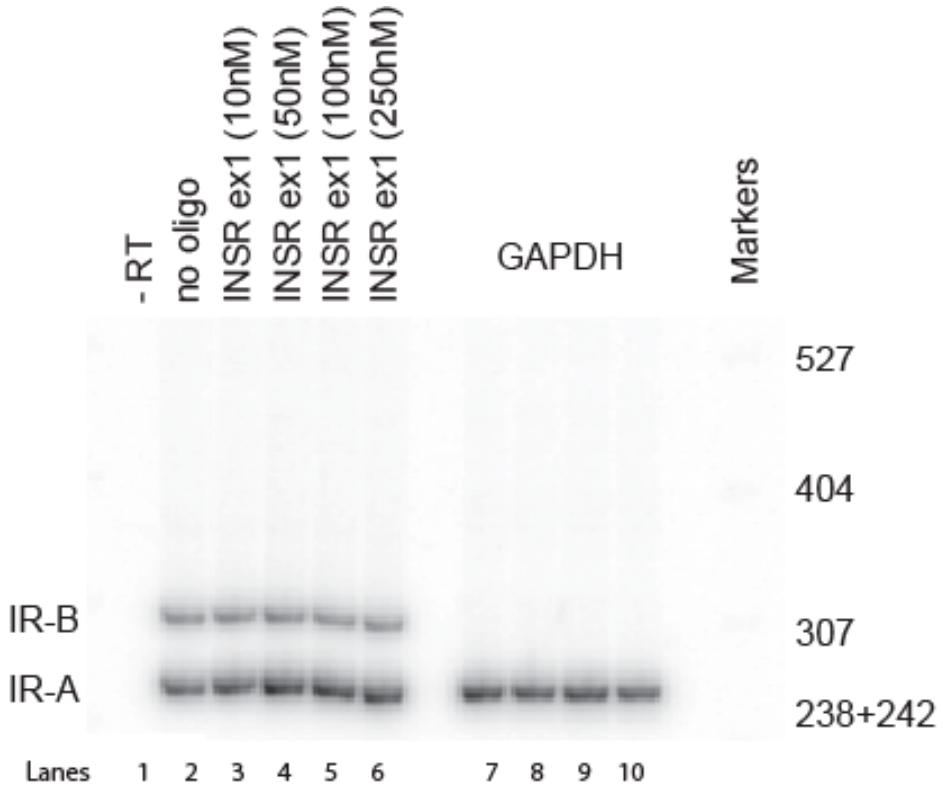


Figure 3.1.3.C Titration of 2′-O-methyl RNA ASO INSR ex1 targeting the IR exon 11. Polyacrylamide gel showing IR isoforms from cells transfected with increasing concentrations of ASO INSR ex1 (10-250 nM). IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion). The right side of the figure represents GAPDH run as a control for mRNA quantity and quality.

Figure 3.1.3.D

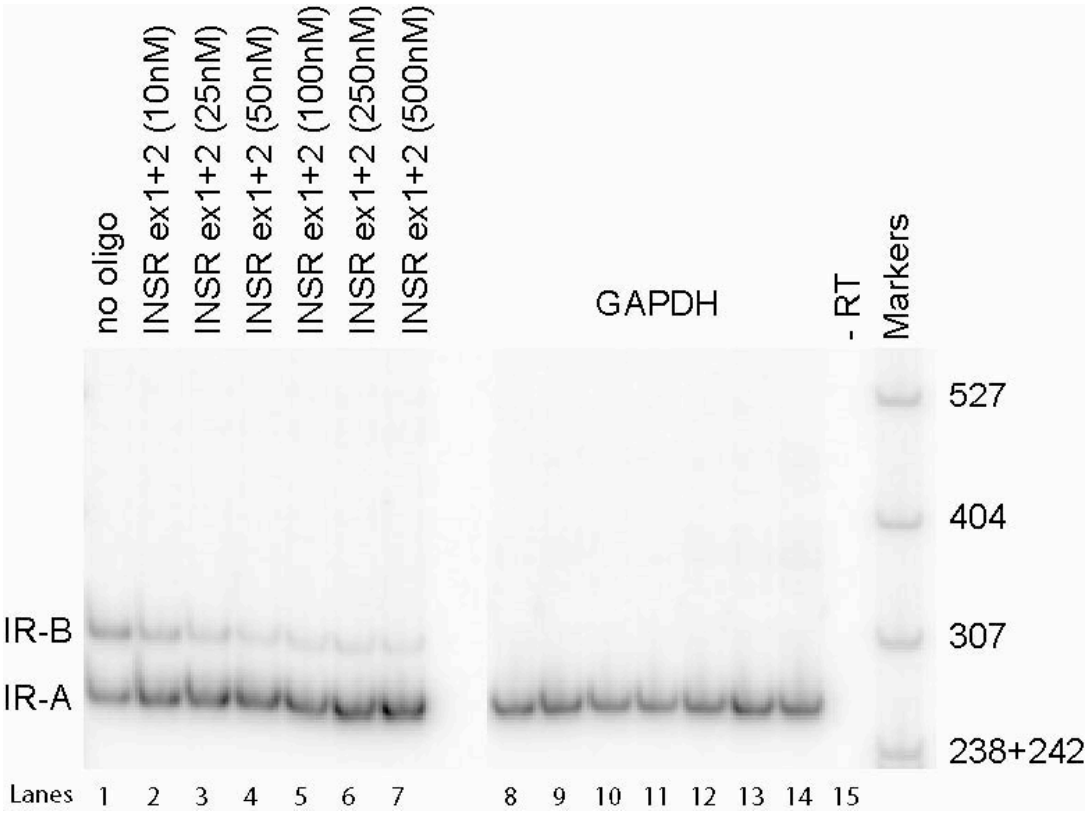


Figure 3.1.3.D Titration of 2′-O-methyl RNA ASO INSR ex1+2 targeting the IR exon 11. Polyacrylamide gel showing IR isoforms from cells transfected with increasing concentrations of ASO INSR ex1+2 (10-500 nM). IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion). The right side of the figure represents GAPDH run as a control for mRNA quantity and quality.

Figure 3.1.3.E

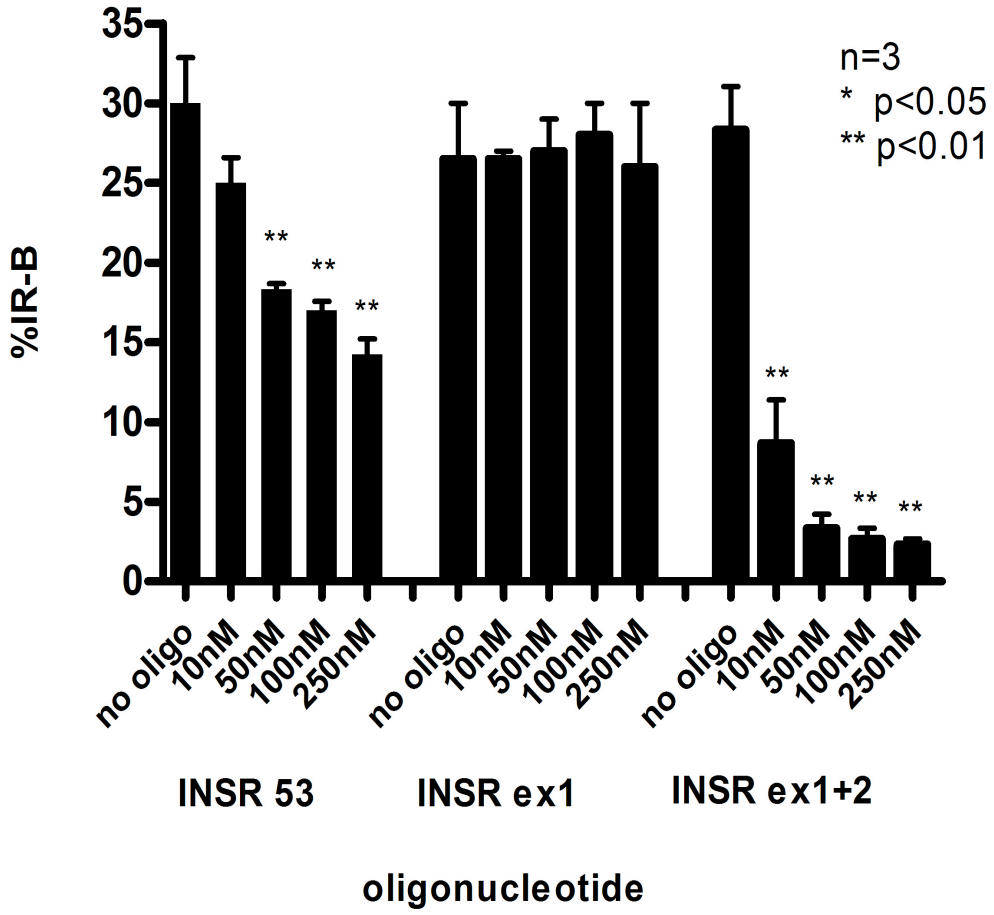


Figure 3.1.3.E Titration of 2′-O-methyl ASOs targeting the IR exon 11.The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection by increasing concentrations (0-250nM) of ASOs INSR 53, INSR ex1 and INSR ex1+2. Quantitative data from 3 replicate experiments. IR exon 11 skipping increased with increasing concentration of ASO INSR 53 and INSR ex1+2. INSR ex1+2 appear more potent than INSR 53 ASO for a given concentration. INSR ex1 had no significant effect on IR exon 11 splicing.

Discussion

This set of experiments investigated the dose-response relationship of ASOs targeting the IR exon 11. The concentration of ASOs transfected into HepG2 cells was between 10 nM to 500 nM. At these concentrations there was no evidence of cellular death, as shown by unchanged GAPDH expression, or upon light microscopy. Cumulative data for the three ASOs assessed is represented graphically in figure 3.1.3.E. With increasing concentrations of effective ASOs INSR 53 (figure 3.1.3.B) and INSR ex1+2 (figure 3.1.3.D), an increase in IR exon 11 skipping was noted. The graph does demonstrate that at 250nM concentration, INSR53 has not reached its maximum effect as the curve has yet to plateau. However for a given concentration, INSR ex1+2 appears more potent than INSR 53. The ASO INSR ex1, previously noted to be ineffective at altering IR exon 11 splicing, was ineffective at all transfection concentrations (figure 3.1.3.C). This confirms the lack of effect of INSR ex1 on IR exon 11 splicing was not a dose-related effect. In keeping with the prediction these results demonstrate that ASOs known to alter the IR exon 11 splicing, do so in a dose-dependent manner. ASOs which are ineffective continue to be ineffective at all doses tested.

3.1.4 Assessment of ASO cellular localisation using fluorescein-tagging.

Background

To modify pre-mRNA splicing, ASOs need to localise in the nuclear cellular compartment. Cationic liposomal transfection agents, such as Lipofectamine 2000 (Gibco/Invitrogen, Paisley, UK), overcome the electrostatic repulsion between the net negatively charged ASO and the negatively charged plasma membrane by forming liposomal complexes. Following cellular internalisation the ASOs, now within an endosomal vesicle, are released from the endosome to traverse to the nucleus. The exact mechanism for this process is unclear.

Differences in cellular localisation between effective ASOs and ineffective ASOs may be an explanation for differences in efficacy. Therefore to further confirm the specificity of effective ASOs it is necessary to demonstrate a similar cellular localisation for both effective and ineffective (scramble) ASOs.

Fluorescein-tagged ASOs allowed direct visualisation of cellular localisation using confocal microscopy. Fluorescein-tagged ASOs may also be used to assess transfection efficiency.

The prediction for this set of experiments was that all equivalent ASOs transfect and localise in a similar manner and that both a fluorescein-tagged ASO and its equivalent fluorescein-tagged scramble would localise in a similar way, with a predominance in the cellular nuclei.

Aim

To identify the cellular localisation of transfected ASOs and confirm a similar cellular distribution between effective and scramble ASOs.

Methods

ASO INSR ex1+2 and its equivalent scramble INSR ex1+2scram were labelled at the 5′ end with fluorescein (Dharmacon, Lafayette, CO, USA). The untagged ASO INSR ex2 was run as a control. ASOs were transfected into HepG2 cells at concentrations 250 nM and incubated for 48 hours at 37^{o}C/5% CO_2. Three hundred microlitres of transfected cell suspension was added to a 6 chamber slide well and incubated separately at 37ºC for 48 hours. The chamber wells were subjected to a series of PBS washes and fixation with 4% paraldehyde/PBS. Three hundred and fifty microlitres of a 2 µg/ml DAPI solution was added to each chamber to visualise the cellular nuclei. DAKO mounting media was applied, cover slips applied and slides were viewed with a fluorescent confocal microscope. Further detailed methodology is given in Chapter 2 - materials and methods. The remaining transfection reactions were processed as before. RNA was harvested, subjected to RT-PCR and run on a polyacrylamide gel to identify the 2 isoforms of the IR and GAPDH as a control for mRNA quantity and quality. Relevant bands were identified and quantified using a Phosphorimager.

Results

Figure 3.1.4.A

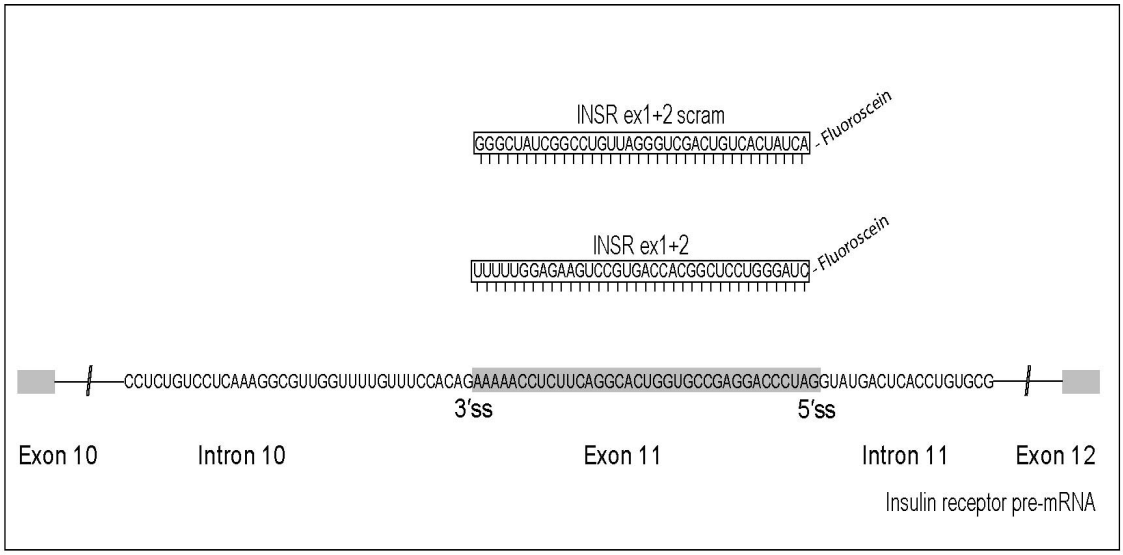


Figure 3.1.4.A Fluorescein tagged ASOs targeting the IR exon 11. Schematic diagram shows a portion of the IR pre-mRNA from exon 10 to exon 12 (not to scale). Exon 11 and surrounding intronic sequences are shown. Exons are delineated by the grey shading, the rest of the sequence being intronic. ASOs are illustrated above the IR pre-mRNA. ASOs INSR ex1+2 and INSR ex1+2 scram were tagged with fluoroscein. Cells were transfected with ASOs at 250 nM. A proportion of the transfection reaction was transferred to a 6 well chamber slide for fluorescent microscopy. RT-PCR was carried out on the total RNA extracted from the remaining transfection reactions.

Figure 3.1.4.B

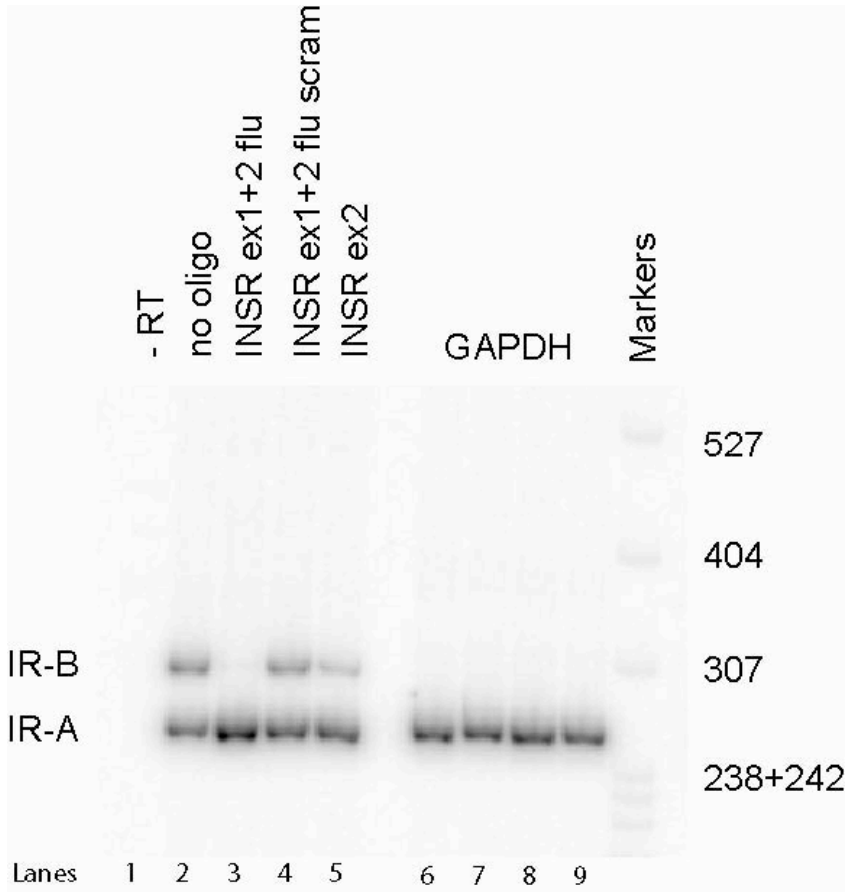


Figure 3.1.4.B Fluorescein-tagged 2′-O-methyl ASOs targeting the IR exon 11. Polyacrylamide gel showing IR isoforms from cells transfected with ASO INSR ex1+2flu and INSR ex1+2flu scram. ASO INSR ex2 was transfected as a positive control. IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion) are shown. GAPDH was run as a control for mRNA quantity and quality.

Figure 3.1.4.C

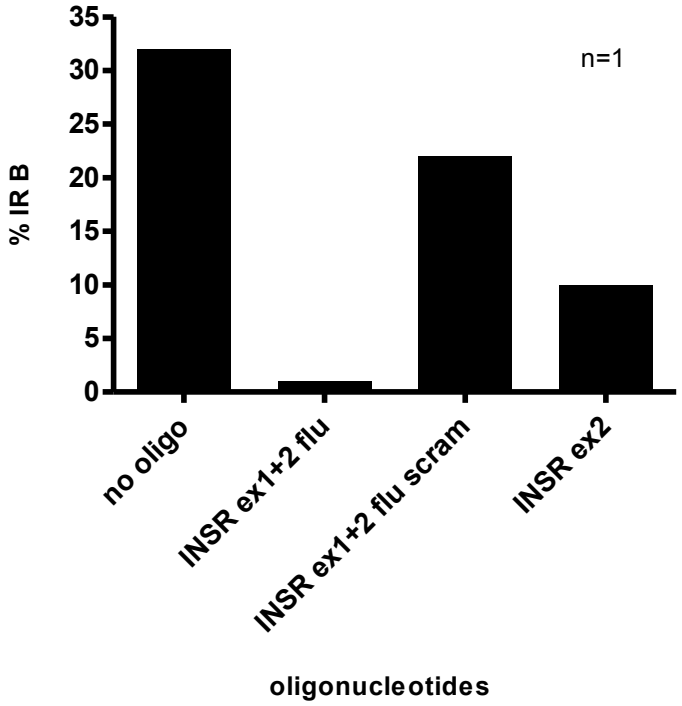


Figure 3.1.4.C Fluorescein-tagged 2′-O-methyl ASOs targeting the IR exon 11. The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection by ASOs INSR ex1+2flu; INSR ex1+2flu scram and INSR ex2 at a concentration of 250nM. Quantitative data from one experiment. ASO INSR ex1+2flu was more effective at causing exon 11 skipping than INSR ex1+2. INSR ex1+2flu scram had no effect on the splicing of IR exon 11.

.

Figure 3.1.4.D

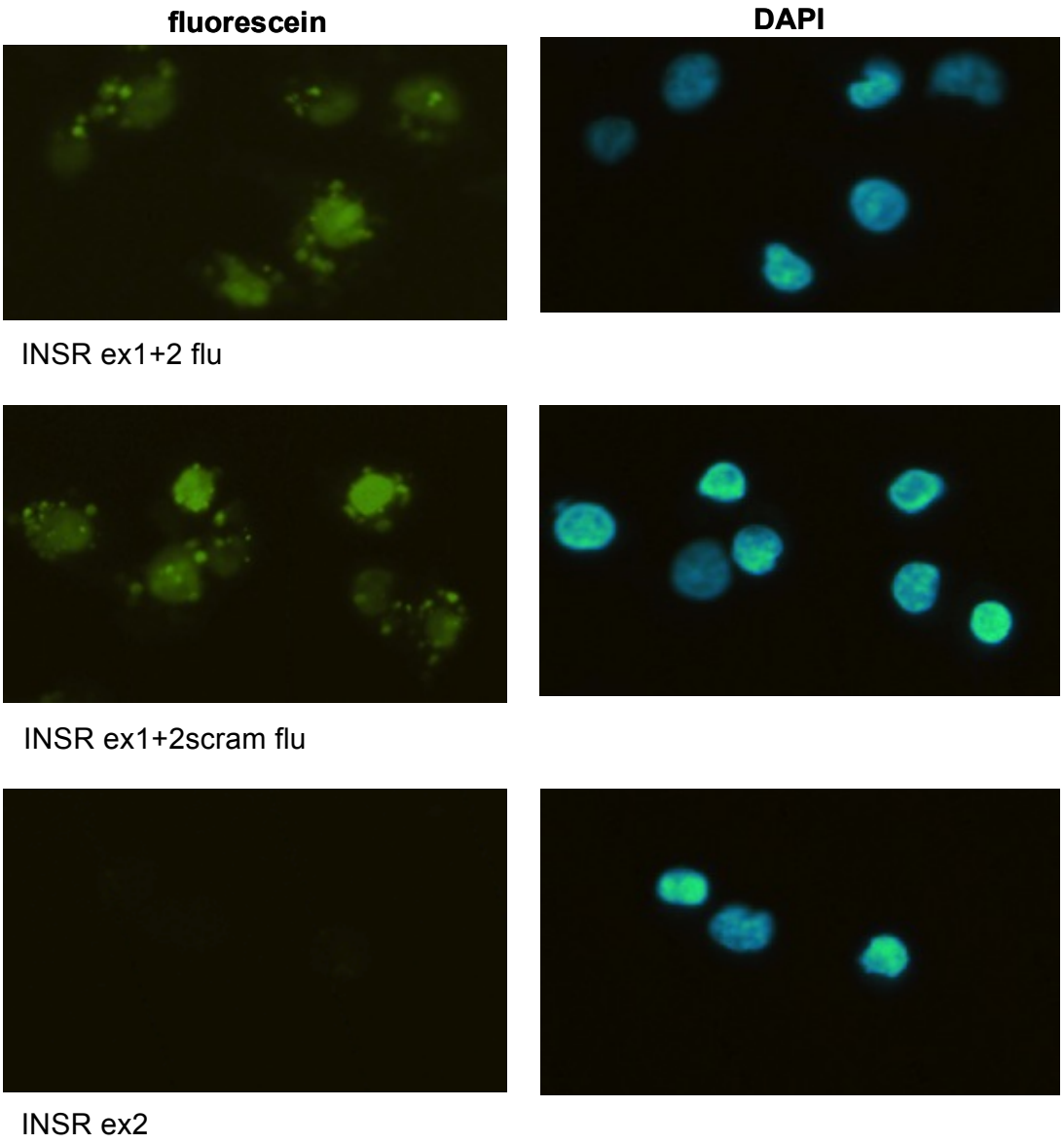


Figure 3.1.4.D Fluorescein-tagged 2′-O-methyl ASOs targeting the IR exon 11. Fluorescein was tagged to ASO INSR ex1+2flu and INSR ex1+2flu scram. DAPI staining identified cellular nuclei (right column). ASOs localised to cytoplasmic inclusions and the nucleus. ASOs and scramble ASOs were distributed equally. INSR ex2, a control without a fluorescein tag showed no luminescence.

Discussion

Figure 3.1.4.B shows a polyacrylamide gel which confirms the efficacy of the ASO INSR ex1+2flu (lane 3) in comparison to the no oligo control (lane 2). The scramble ASO INSR ex1+2flu scram (lane 4) appears to have a minimal effect on the splicing of the IR exon 11 but in the context of an experiment not repeated in triplicate (figure 3.1.4.C). The fluorescein-tagged ASO INSR ex1+2flu was more effective at causing IR exon 11 skipping (IR-B <1%) than the untagged ASO INSR ex1+2 (IR-B 4%) assessed in Chapter 3.1.1. This difference may be ascribed to the fact that the fluorescein-tagged ASOs were highly purified whereas the untagged ASOs were not. Fluorescein-tagged ASOs may therefore have appeared more potent at a given concentration than the non-tagged equivalent. An alternative explanation is that some inherent property of the fluorescein tag increases transfection efficiency, or that the fluorescein increases the exon-skipping effect of the ASO.

Figure 3.1.4.D demonstrates that the fluorescein-tagged ASO INSR ex1+2flu was localised both in the cytoplasm and more intensely in the nuclei (figure 3.1.4.D - top row). Scramble ASO INSR ex1+2flu scram localised in an identical pattern (figure 3.1.4.D - middle row) despite being ineffective at altering IR exon 11 splicing. Difference in efficacy between effective ASOs and ineffective ASOs are therefore not related to differences in cellular localisation.

This set of experiments confirms the prediction that the effective exon-skipping ASOs localise in identical patterns, with predominance in nuclei, to those which are ineffective in causing exon skipping. ASO localisation is independent of the effects these ASOs have on IR exon 11 splicing.

3.1.5 TOES ASOs targeting the IR exon 11.

Background

Previous experiments from this chapter had shown the ability of ASOs to cause IR exon 11 skipping. No ASO tested to this point had demonstrated the ability to increase exon 11 inclusion. To achieve this, an adjunct to the current approach was the addition of 'non-complimentary' tail sequences designed to enhance the splicing process by various means. 'Tailed oligonucleotide enhancement of splicing' or TOES ASOs consist of an ASO complementary to a chosen sequence and a non-complementary tail mimicking an ESE sequence, designed with the intention of recruiting enhancing SR proteins. This approach has been used successfully both in vitro and in vivo (Owen et al. 2011; Skordis et al. 2003). SMN1 defects are known to cause the disease spinal muscular atrophy. The paralogous gene, SMN2 encodes transcripts which do not include exon 7 leading to an unstable SMN protein (Khoo et al. 2009). In the above studies TOES ASOs were designed to bind to SMN2 pre-mRNA and to recruit enhancing splicing factors (ASF/SF2) to exon 7 (Owen et al. 2011; Skordis et al. 2003). This approach was able to successfully increase exon 7 inclusion, leading to restoration of SMN protein function in cells derived from patients with spinal muscular atrophy. It was therefore reasoned that a similar approach might be able to increase the inclusion of IR exon 11.

Aim

To increase the IR exon 11 inclusion with a trans-splicing approach using TOES ASOs.

Methods

TOES ASOs were designed with a 2′-O-methyl RNA ASO base as used in earlier experiments (figure 3.1.5.A). Sequences are given in Chapter 6.1 - appendix. The ASO was targeted to the 5′ end of the IR exon 11 on the basis of experiments from Chapter 3.1.1. This portion of exon 11 appeared to have a neutral effect on the splicing of the exon when targeted and would therefore not be expected to inhibit exon 11 inclusion. The ESE tail sequence was identical to that used in the rescue of exon 7 in SMN2 (Skordis et al. 2003). Tails were added to either to the 3′ or 5′ end of the ASO. An ESE tail, unattached to an ASO, was transfected as a control.

TOES ASOs were reverse-transfected into HepG2 cells using Lipofectamine 2000 (Gibco/Invitrogen, Paisley, UK) at varying concentration according to the standard protocol provided in Chapter 2 - materials and methods. ASO INSR ex1+2 was used as a positive control. HepG2 cells were incubated 48 hours under standard conditions. RNA was subsequently harvested and subjected to RT-PCR to identify the 2 IR isoforms and GAPDH as a control for mRNA quality and quantity. Products were run on a polyacrylamide gel and the radioactivity from each sample quantified using a Phosphorimager. Experiments were repeated in triplicate.

Results

Figure 3.1.5.A

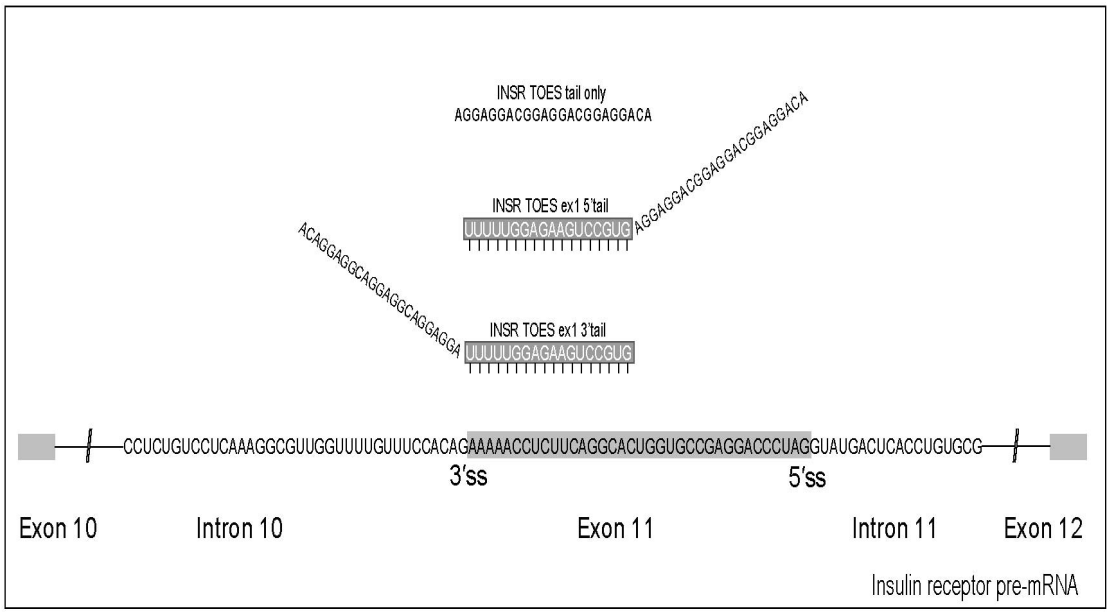


Figure 3.1.5.A TOES ASOs targeting the IR exon 11. Schematic diagram shows a portion of the IR pre-mRNA from exon 10 to exon 12 (not to scale). Exon 11 and surrounding intronic sequences are shown. Exons are delineated by the grey shading and black characters, the rest of the sequence being intronic. TOES ASOs are illustrated above the IR pre-mRNA with grey shadowing and white characters. TOES ASOs, 18 nucleotides long, were designed to block the 5′ half of exon 11. A tail, 23 nucleotides long, simulating an exonic splicing enhancer sequence (ESE), was attached either to the 3′ or the 5′ end of the ASO designated INSR TOES ex1 3′ tail or INSR TOES ex1 5′ tail respectively. The ESE tail sequence, unattached to a 2′-O-methyl RNA sequence, designated INSR TOES tail only, was transfected into cells as a control. Cells were transfected with ASOs at 250 nM concentration. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.

Figure 3.1.5.B

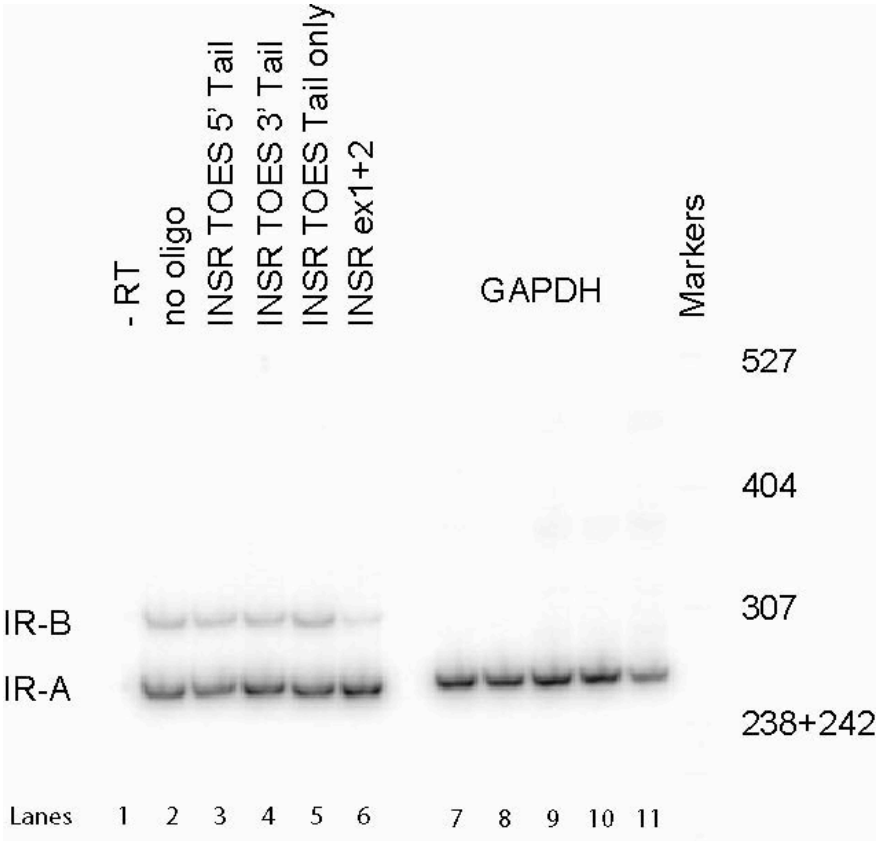


Figure 3.1.5.B TOES ASOs targeting the IR exon 11. Polyacrylamide gel from a representative experiment showing IR isoforms from HepG2 cells transfected with ASOs at 250 nM concentration. Each column represents a well of transfection with ASOs marked above the column. The left side of the figure represents the two IR isoforms - IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion). The right side of the figure represents GAPDH which was run as a control for mRNA quantity and quality.

Figure 3.1.5.C

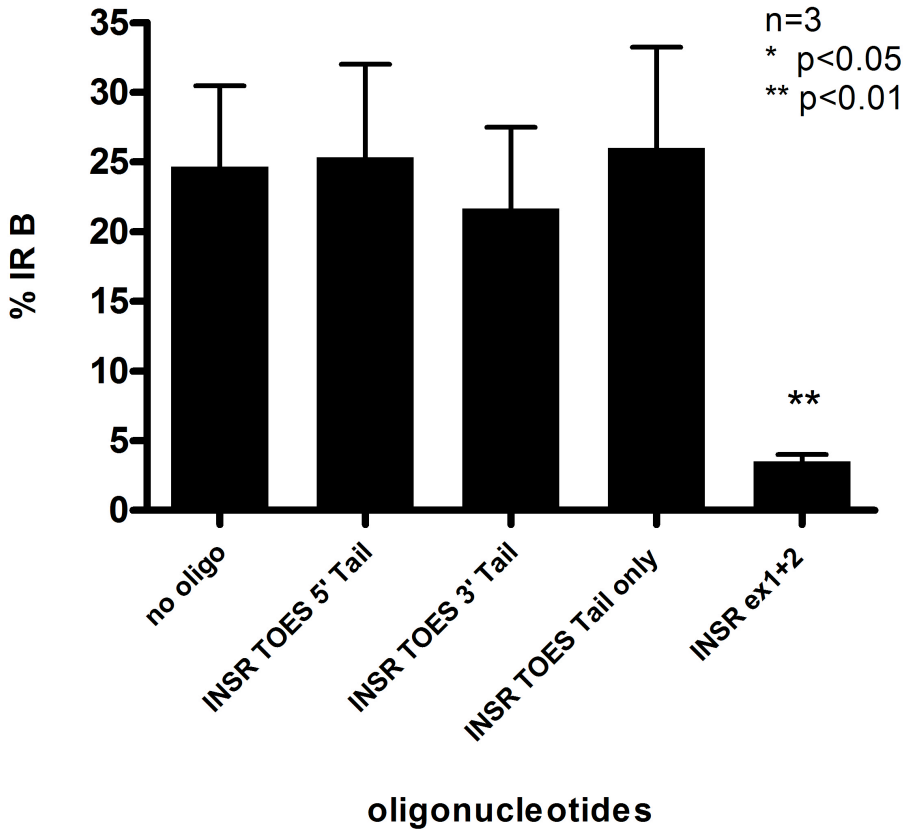


Figure 3.1.5.C TOES ASOs targeting the IR exon 11. The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection by TOES molecules at a concentration of 250nM. Quantitative data from three independent replications of the experiment (error bars show the S.E.M.). Overall one-way repeated measures ANOVA of the dataset showed $p<0.0001$, indicating statistically significant differences between groups. Dunnett's multiple comparison test was used to compare the no oligo control versus TOES transfected cells. The TOES molecules had no effect on splicing in contrast to the ASO INSR ex1+2.

Discussion

The representative gel in figure 3.1.5.B demonstrates that the TOES ASOs designed to target the IR exon 11 do not enhance exon inclusion (compare lane 2 with lanes 3 and 4). No effect was noted with increasing concentration of TOES ASOs between 250-1000 nM (data not shown). The positive control, ASO INSR ex1+2 (lane 6) continued to cause IR exon 11 skipping, confirming transfection had occurred (figure 3.1.5.C). The position of the tail, at the 3′ or 5′ end of the ASO, did not influence the degree of exon 11 inclusion. The ESE tail transfected alone (lane 5) did not alter splicing. A TOES ASO targeting the 3′ end of the IR exon 11 was not studied.

A previous study noted a TOES ASO increased inclusion of exon 7 in the SMN2 transcript in vivo from 57% inclusion up to 84% inclusion. This affect was initially noted at a 100 nM concentration (Skordis et al. 2003). The aim, to increase IR exon 11 inclusion using a tailed-oligonucleotide localising an additional enhancer sequence, was not achieved. This is discussed further in Chapter 4.2.

Chapter 3.2

PNA and PNA-peptide hybrid antisense oligonucleotides targeting the insulin receptor exon 11

3.2.1 PNA ASO targeting the IR exon 11

Experiments in Chapter 3.1 demonstrated that 2′-O-methyl RNA ASOs targeting key intronic and exonic sites around the IR exon 11 were unable to stimulate exon 11 inclusion. 18mer 2′-O-methyl RNA ASOs may have been too long to target silencer regions in isolation within exon 11, which may have facilitated exon inclusion. An alternative strategy was employed to target shorter regions within exon 11. This required the use of different oligonucleotide chemistry. Peptide nucleic acids are synthetic DNA-like compounds which substitute the DNA phosphodiester backbone for a pseudopeptide backbone ((N-(2-aminoethyl)glycine). A methylene carbonyl link binds the nucleotide bases to this. PNA ASOs have several advantages over earlier oligonucleotides chemistries. These included increased stability due to the non-charged peptide backbone as well as resistance to breakdown by nucleases and peptidase. They also demonstrate increased binding specificity and affinity for their targets (Egholm et al. 1993; Jensen et al. 1997). The increased stability that PNA ASOs offered allowed smaller targets within exon 11 to be pursued. The chemistry was also chosen as it formed the basis of PNA-peptide hybrid molecules which were to be used to attempt exon 11 inclusion in subsequent work (Chapter 3.3).

The prediction was that targeting shorter regions of the IR exon 11 would target enhancer or silencer regions more precisely. Shorter (12mer) PNA ASOs targeting the 5′, 3′ and central regions of the IR exon 11 would cause either IR exon 11 inclusion or skipping.

Aim

To assess the effects of PNA ASOs targeting the IR exon 11.

Methods

PNA ASOs, 12 nucleotides in length were designed to target the 36 nucleotides of the IR exon 11. The PNA targeting the 12 nucleotides at the 5′ end of exon 11 was designated INSR PNA ex4, the PNA ASO targeting the middle 12 nucleotides was designated INSR PNA ex5 and the 12 nucleotides at the 3′ end was designated INSR PNA ex6. PNA ASOs were produce by Bio-Synthesis Inc (Texas, U.S.A).

PNA ASOs were reverse-transfected into HepG2 cells for 48 hours as per the previously detailed methodology (Chapter 2 - materials and methods). PNA ASOs were transfected at a concentration of 250 nM initially. The 2′-O-methyl RNA ASO, INSR ex2 was transfected as a positive control. Following RNA harvest, RT-PCR was carried out with primers identifying the 2 isoforms of the IR and GAPDH. Products were then run on a polyacrylamide gel.

Figure 3.2.1.A

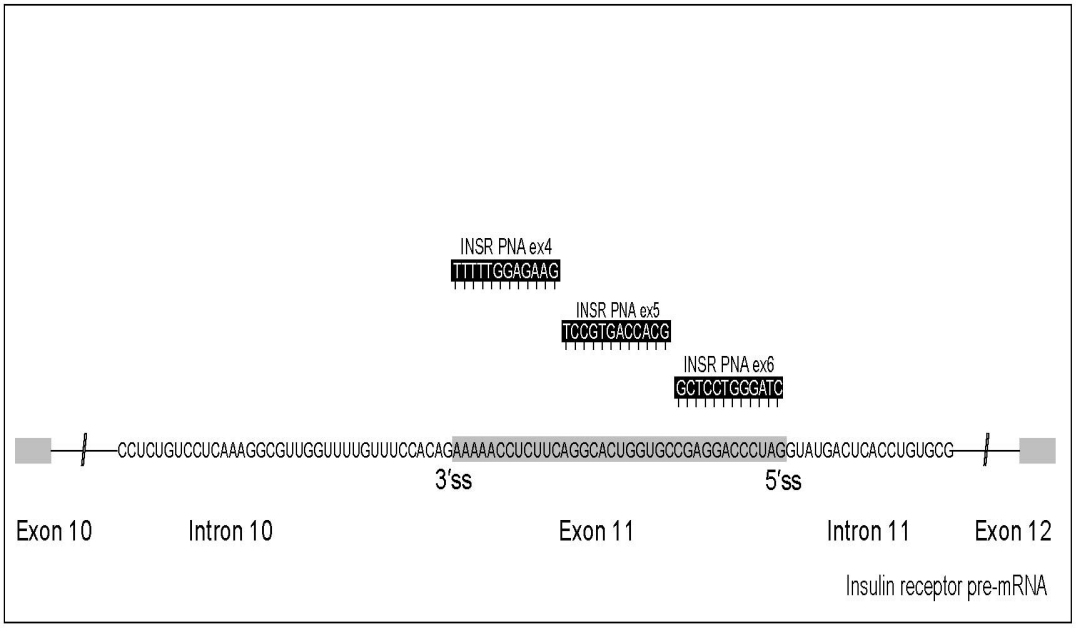


Figure 3.2.1.A PNA ASOs targeting the IR exon 11. Schematic diagram shows a portion of the IR pre-mRNA from exon 10 to exon 12 (not to scale). Exon 11 and surrounding intronic sequences are shown. Exons are delineated by the grey shading, the rest of the sequence being intronic. PNA ASOs are illustrated above the IR pre-mRNA. PNA ASOs, 12 nucleotides long, were designed to block a third each of exon 11. INSR PNA ex4 covered the 5′ end of exon 11, INSR PNA ex5, the middle 12 nucleotides and INSR PNA ex6 covered the 3′ end of exon 11. Cells were transfected with ASOs at 250 nM concentration. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.

Figure 3.2.1.B

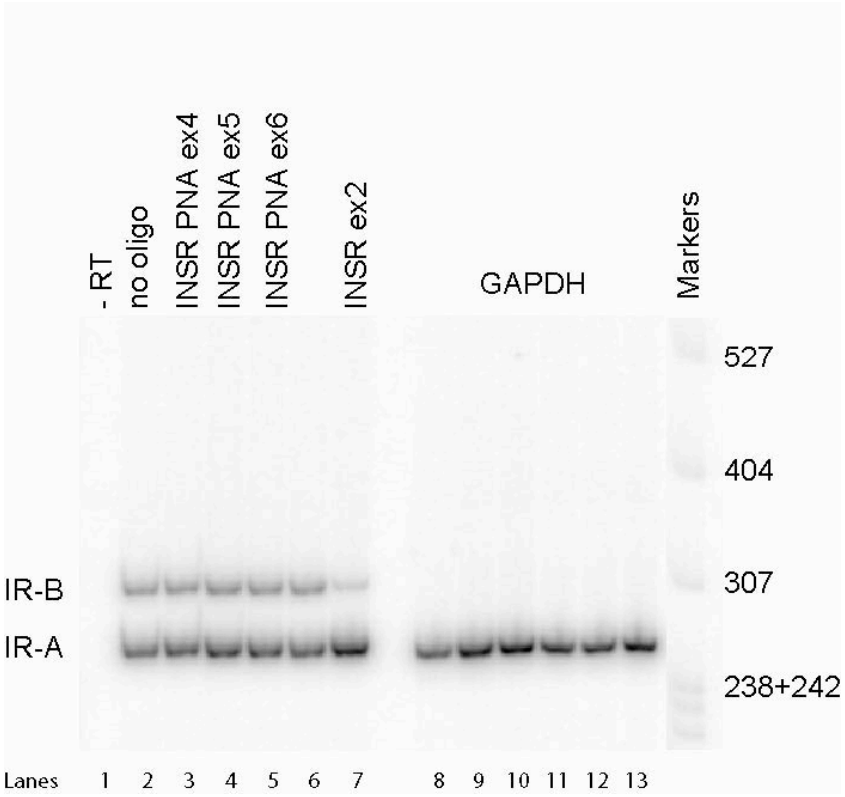


Figure 3.2.1.B PNA ASOs targeting the IR exon 11. Polyacrylamide gel showing IR isoforms from cells transfected with PNA molecules at 250 nM; INSR PNA ex4; INSR PNA ex5 and INSR PNA ex6. INSR ex2 was run as a positive control. IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion) are shown on the left side of the figure. GAPDH was run as a control for mRNA quantity and quality.

Figure 3.2.1.C

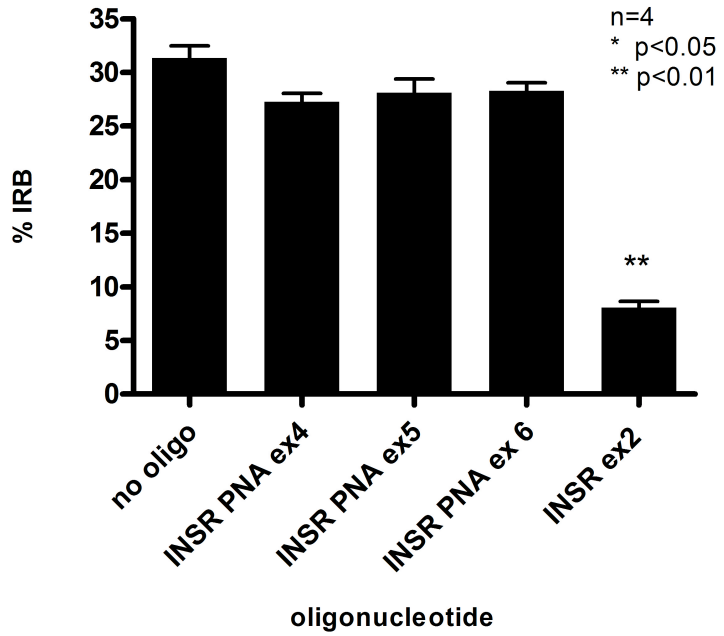


Figure 3.2.1.C PNA ASOs targeting the IR exon 11. The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection by PNA ASOs at a concentration of 250nM. Quantitative data from three independent replications of the experiment (error bars show the S.E.M.). Overall one-way repeated measures ANOVA of the dataset showed $p<0.0001$, indicating statistically significant differences between groups. Dunnett's multiple comparison test was used to compare the no-oligo control versus PNA molecule transfected cells. The PNA molecules were not effective at altering IR exon 11 splicing. The 2′-O-methyl RNA ASO INSR ex2 was the only effective ASO.

Discussion

Figure 3.2.1.B shows a representative gel following transfection of INSR PNA ex4, ex5 and ex6 into HepG2 cells at a 250 nM concentration for 48 hours and the graph in figure 3.2.1.C represents the results of experiments completed in triplicate. These show that PNA ASOs targeting consecutive 12 nucleotide stretches of the IR exon 11 have no effect on the splicing (compare no oligo in lane 2 with PNA INSR ex4, ex5 and ex6 in lanes 3, 4 and 5 respectively). The positive control INSR ex2 maintains its effect on IR exon 11 splicing. These results are against the prediction that 12mer PNA ASOs would cause either IR exon inclusion or skipping. They do not support the prediction that PNA ASOs can more precisely target enhancer or silencer regions within the exon. Several explanations may be postulated for these findings: firstly, inadequate concentration of PNA ASO; secondly, poor transfection efficiency results in failure to bind the appropriate sites; thirdly, failure to block key splicing interaction e.g. with the spliceosome and finally, simultaneous blocking of competing enhancer and silencer sites in exon 11.

3.2.2 PNA ASOs targeting an equivalent sequence to ASO INSR ex2 in the IR exon 11 - Are the effects equivalent?

Background

Data from Chapter 3.2.1 has shown that PNA ASOs targeting the IR exon 11 do not alter its splicing. This was an unexpected finding given that some of the 2′-O-methyl RNA ASOs targeting exon 11 did alter splicing. The PNA ASOs tested however were not directly comparable to the 2′-O-methyl RNA ASOs with respect to the sequences they targeted. It was therefore necessary to confirm whether PNA ASOs and equivalent 2′-O-methyl RNA ASOs have the same effects on splicing.

The specific prediction was that PNA ASOs and 2′-O-methyl RNA ASOs had equivalent effects on splicing and that a PNA ASO targeting the same sequence in the IR exon 11 as the 2′-O-methyl ASO INSR ex2 would cause exon 11 skipping.

Aim

To confirm the similar effects of PNA ASOs and 2′-O-methyl ASO targeting the same sequence on IR exon 11 splicing.

Methods

A PNA ASOs, designated PNA INSR ex2 (18mer) was designed to target the same sequence as the 2′-O-methyl ASO INSR ex2 (figure 3.2.2.A and Chapter 6.1 - appendix). PNA INSR ex2 (18mer) was specifically designated ‘18mer’ to avoid confusion with another PNA ASO designated PNA INSR ex2. PNA INSR ex2 formed the basis for a PNA-peptide hybrid molecule tested subsequently which targeted the same sequence as PNA INSR ex2 (18mer) except the 3 nucleotides at the 3′ end of exon 11 and was therefore a 15 mer. A scramble sequence of PNA

INSR ex2 (18mer) was constructed as a further control and designated PNA INSR ex2scram (18mer).

PNA INSR ex2 (18mer) and PNA INSR ex2scram (18mer) were transfected at a concentration of 250 nM in HepG2 cells for 48 hours according to the standard protocol as detailed in Chapter 2 - materials and methods. The 2′-O-methyl ASO INSR ex2 was transfected as a positive control. Following RNA harvest and RT-PCR using primers to identify the IR isoform and GAPDH, products were run on a polyacrylamide gel.

Figure 3.2.2.A

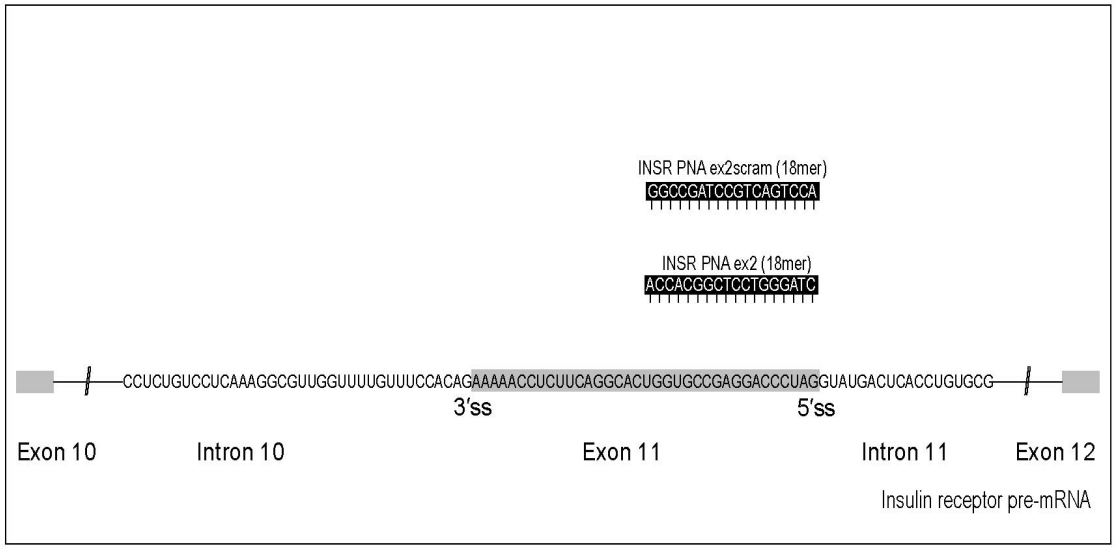


Figure 3.2.2.A PNA ASOs targeting an equivalent sequence to ASO INSR ex2 in the IR exon 11. Schematic diagram shows a portion of the IR pre-mRNA from exon 10 to exon 12 (not to scale). Exon 11 and surrounding intronic sequences are shown. Exons are delineated by the grey shading, the rest of the sequence being intronic. PNA ASOs are illustrated above the IR pre-mRNA. PNA ASOs, 18 nucleotides long, were designed to block the 3′ half of exon 11, designated INSR PNA ex2 (18mer). A scramble control sequence designated INSR PNA ex2scram (18mer) was also designed. Cells were transfected with ASOs at 250 nM concentration. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.

Figure 3.2.2.B

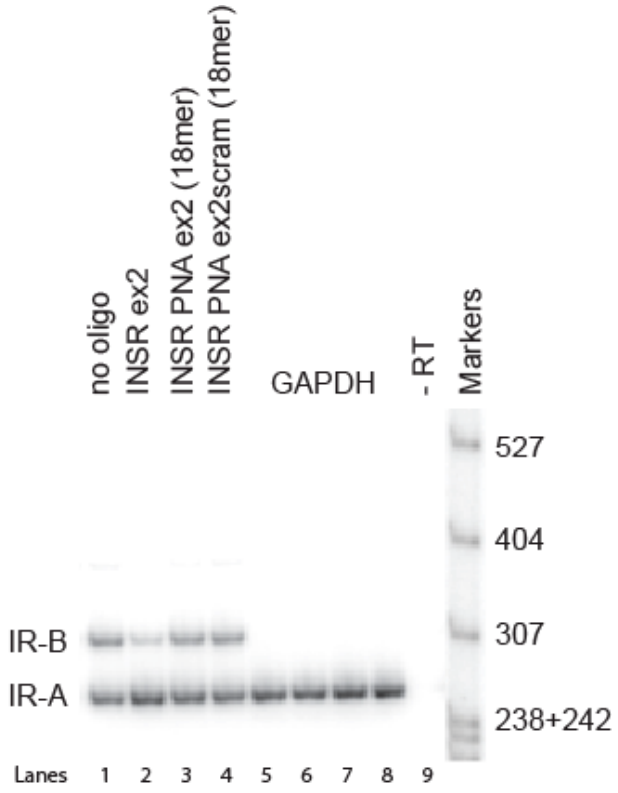


Figure 3.2.2.B PNA ASOs targeting an equivalent sequence to ASO INSR ex2 in the IR exon 11. Polyacrylamide gel from a representative experiment showing IR isoforms from HepG2 cells transfected with PNA ASOs. Each column represents a well of transfection with the ASO marked above the column. The left side of the figure represents the two IR isoforms - IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion). 2′-O-methyl RNA ASO INSR ex2 was run as a positive control. GAPDH was run as a control for mRNA quantity and quality.

Figure 3.2.2.C

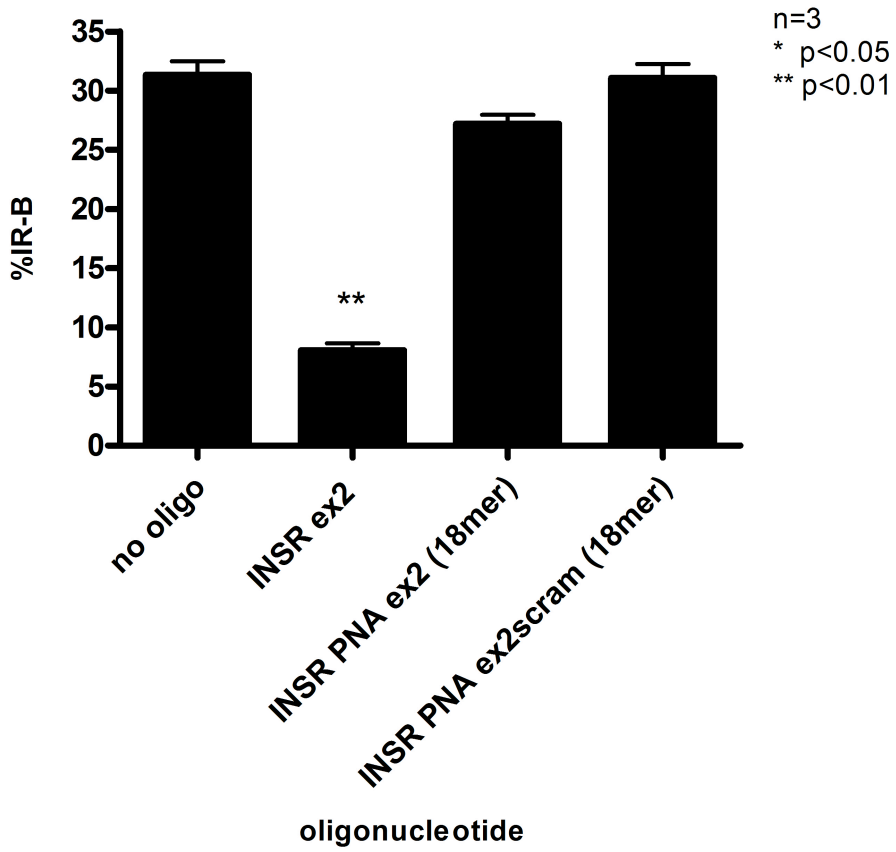


Figure 3.2.2.C PNA ASOs targeting an equivalent sequence to ASO INSR ex2 in the IR exon 11. The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection by ASOs in HepG2 cells at a concentration of 250 nM. Quantitative data from three independent replications of the experiment (error bars show the S.E.M.). Overall one-way repeated measures ANOVA of the dataset showed $p<0.0001$, indicating statistically significant differences between groups. Dunnett's multiple comparison test was used to compare the no-oligo control versus ASO transfected cells: statistically significant differences were only found between the no oligo control transfected cells and the positive control ASO INSR ex2.

Discussion

Figure 3.2.2.B shows a representative gel from experiments involving the transfection of PNA INSR ex2 (18mer) and PNA INSR ex2scram (18mer). Graph 3.2.2.C shows the cumulative results of 3 replicate experiments. At a concentration of 250 nM, neither PNA INSR ex2 (18mer) (figure 3.2.2.B lane 3) or PNA INSR ex2scram (18mer) (lane 4) had any significant effect on IR exon 11 splicing in comparison to the no oligo transfected cells (lane 1). This is also in comparison to the positive control INSR ex2 (lane 2) which caused significant exon 11 skipping. Other PNA molecules had been transfected up to a concentration of 1 μM without demonstrating an effect on splicing. (Chapter 3.2.3 and 3.2.4).

These findings contradict the prediction that equivalent PNA ASOs and 2′-O-methyl RNA ASOs would have similar effect on splicing of the IR exon 11. Despite targeting the same sequences, assumptions can not be made about the effects of ASOs on splicing with different chemistries.

3.2.3 PNA-peptide hybrid ASOs targeting the 5′ end of the IR exon 11.

Background

Results from Chapters 3.1 and 3.2 have shown that neither 2′-O-methyl RNA ASOs, TOES ASOs or PNA ASOs are able to increase IR exon 11 inclusion. An alternative strategy to achieve this was the use of PNA-peptide hybrid ASOs to enhance exon 11 splicing. PNA-peptide hybrids ASOs or 'ESSENCE' (exon specific splicing enhancement by small chimeric effectors) are ASOs with arginine/serine tails simulating the RS domain of splicing enhancers. These compounds have been used in various models to correct aberrant splicing including the rescue of SMN2 exon 7, BRCA exon and Bcl-2 (Cartegni et al. 2003; Wilusz et al. 2005). Exon inclusion is mediated in two ways. The PNA ASO may bind and block a cis-acting inhibitory site resulting in enhanced splicing. Secondly a tail consisting of ten arginine serine repeats, mimicking the RS domain of an SR protein, tails the antisense portion of the molecule and should recruit relevant enhancing splicing factors and the spliceosome to the exon thereby enhancing splicing. This approach has been adapted in this study to enhance the splicing of the IR exon 11.

The specific prediction for these experiments was that modified ASOs (PNA-peptide hybrid ASOs) with the ability to recruit and interact with splicing factors and/or the spliceosome would enhance exon 11 splicing.

Aim

To assess the ability of PNA-peptide hybrid ASOs to enhance exon 11 inclusion by targeting its 5′ end.

Methods

In the published studies the PNA ASO targets inhibitory sites resulting in the enhancement of splicing even in the absence of an RS tail (Cartegni et al. 2003). No inhibitory sites had been identified within exon 11 so a neutral region was targeted instead. Chapter 3.1 demonstrated that ASO INSR ex1 targeting the 5′ half of exon 11 had no effect on splicing. This region was therefore targeted with 15mer PNA ASO targeting nucleotides 4 - 18 of exon 11. The first 3 nucleotides of exon 11 were not targeted on the basis that this may hinder the 3′ splice site recognition by the spliceosome and hinder the splicing process. The arginine/serine tail (10 repeats) was attached to the PNA ASO via a glycine-glycine linker. Two PNA-peptide hybrid ASOs were constructed with the RS domain tail at either the 3′ end of the molecule or the 5′ end (figure 3.2.3.A) designated INSR PNA ex1 3′ tail and INSR PNA ex1 5′ tail respectively. Additional compounds where constructed to serve as controls and these included the PNA ASO alone, designated INSR PNA ex1 and the RS domain tail with the glycine-glycine linker alone, designated INSR PNA tail only. PNA-peptide hybrids were manufactured by Bio-Synthesis (Texas, U.S.A).

PNA-peptide hybrid ASOs were reverse-transfected into HepG2 cells at various concentrations for 48 hours as per the detailed methodology given in Chapter 2 - materials and methods. The 2′-O-methyl RNA ASO, INSR ex2 was transfected as a positive control along with the other controls. RNA was subsequently harvested and subjected to RT-PCR. PCR primers identified the 2 isoforms of the IR and GAPDH

which was run as a control for mRNA quantity and quality. Products were run on a polyacrylamide gel.

Results

Figure 3.2.3.A

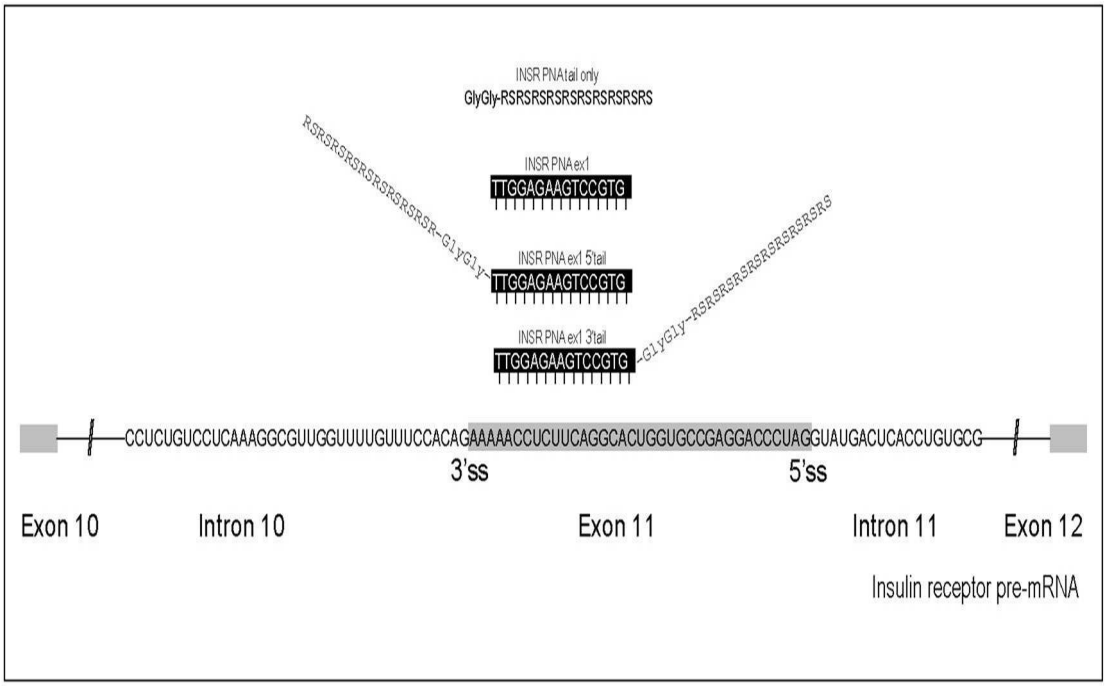


Figure 3.2.3.A PNA-peptide hybrid ASOs targeting the 5′ end of IR exon 11. Schematic diagram shows a portion of the IR pre-mRNA from exon 10 to exon 12 (not to scale). Exon 11 and surrounding intronic sequences are shown. Exons are delineated by the grey shading and black characters, the rest of the sequence being intronic. PNA-peptide hybrid ASOs are illustrated above the IR pre-mRNA with black shadowing and white characters. The PNA backbone was 15 nucleotides long binding the 5′ end of exon 11 after the 3 nucleotides next to the intron/exon border. A tail with 10 arginine serine (RS) repeats connected by a glycine-glycine link was attached either to the 3′ end of the PNA, designated INSR PNA ex1 3′ tail or the 5′ end of the PNA designated INSR PNA ex1 5′ tail. Both the arginine/serine tail alone (INSR PNA tail only) and the PNA alone (INSR PNA ex1) were transfected separately as controls. Cells were transfected with ASOs at varying concentrations. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.

Figure 3.2.3.B

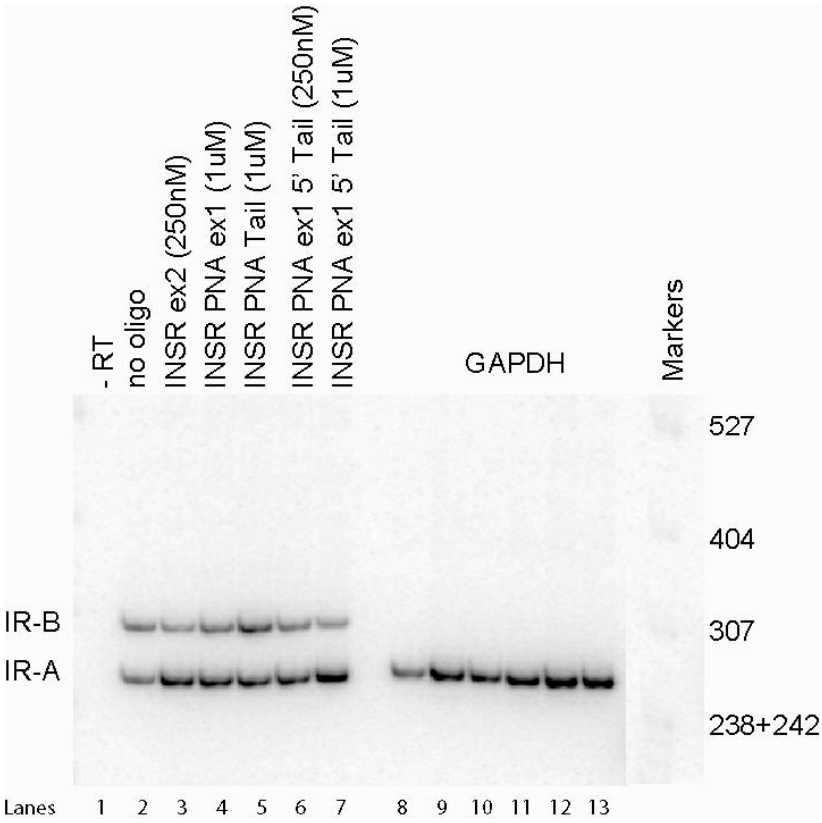


Figure 3.2.3.B PNA-peptide hybrid ASOs (5′ tail) targeting the 5′ end of IR exon 11. Polyacrylamide gel showing IR isoforms from cells transfected with INSR PNA ex1 with peptide tails at varying concentrations. ASO INSR ex2 was used as a positive control. IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion) are shown. GAPDH was run as a control for mRNA quantity and quality.

Figure 3.2.3.C

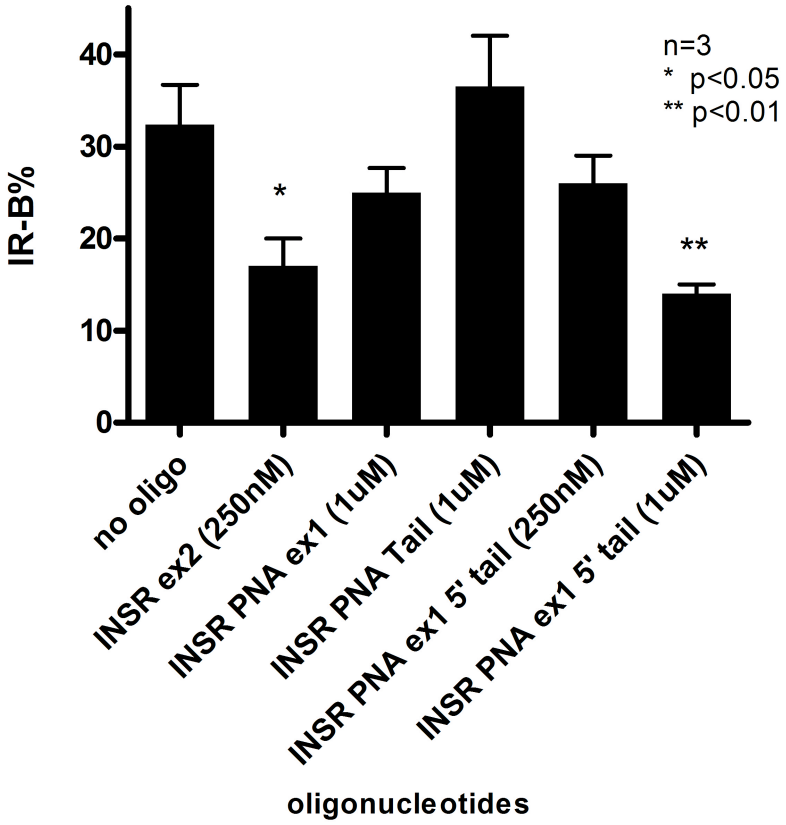


Figure 3.2.3.C PNA-peptide hybrid ASOs (5′ tail) targeting the 5′ end of IR exon 11. The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection at varying concentrations (0-1μM). Quantitative data from three independent replications of the experiment (error bars show the S.E.M.). Overall one-way repeated measures ANOVA of the dataset showed $p<0.0001$, indicating statistically significant differences between groups. Dunnett's multiple comparison test was used to compare the no oligo control versus PNA ASO transfected cells. The PNA - peptide hybrid ASO (5′ tail) showed a statistically significant reduction in the proportion of IR-B at a 1uM concentration.

Figure 3.2.3.D

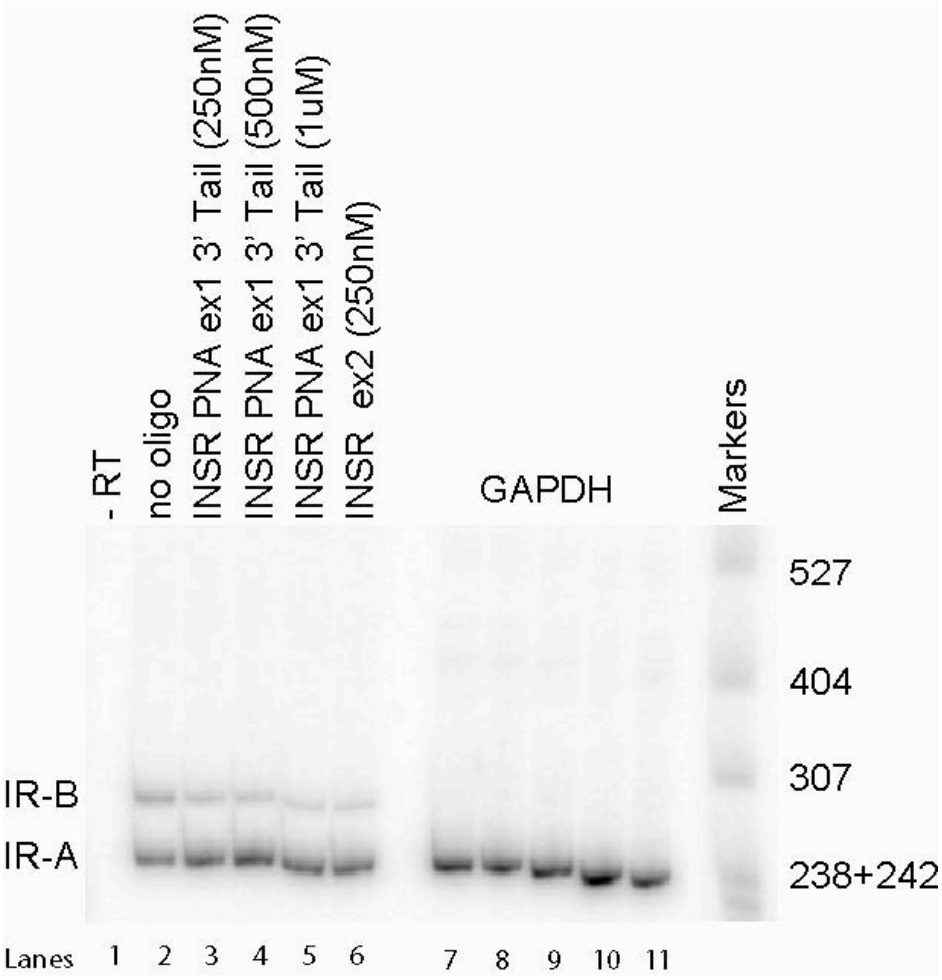


Figure 3.2.3.D PNA-peptide hybrid ASOs (3′ tail) targeting the 5′ end of IR exon 11. Polyacrylamide gel showing IR isoforms from cells transfected with INSR PNA ex1 3′ tails at varying concentrations. ASO INSR ex2 was used as a positive control. IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion) are shown. GAPDH was run as a control for mRNA quantity and quality.

Figure 3.2.3.E

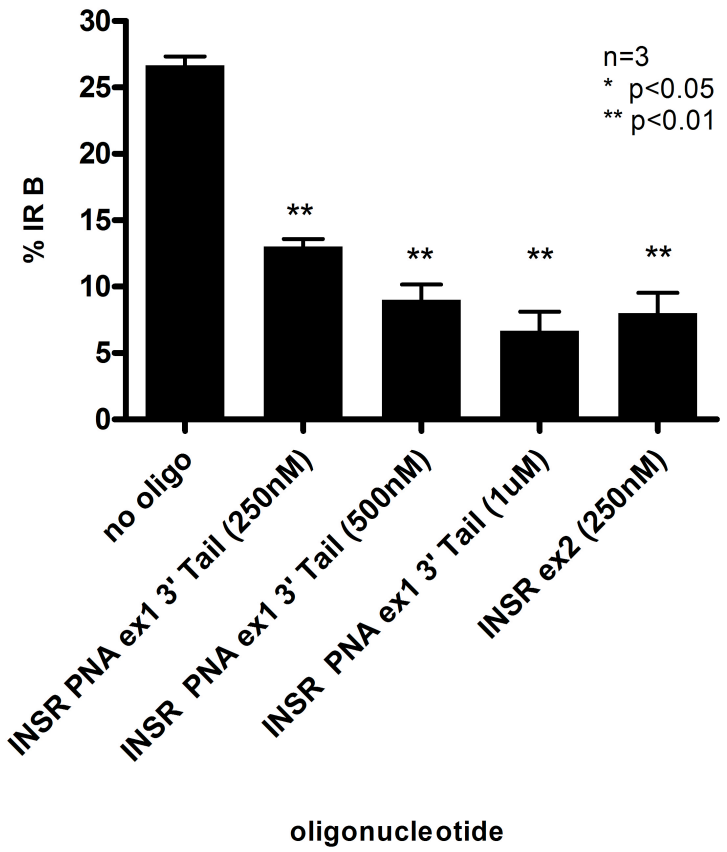


Figure 3.2.3.E PNA-peptide hybrid ASOs (3′ tail) targeting the 5′ end of IR exon 11. The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection at varying concentrations (0-1μM). Quantitative data from three independent replications of the experiment (error bars show the S.E.M.). Overall one-way repeated measures ANOVA of the dataset showed $p<0.0001$, indicating statistically significant differences between groups. Dunnett's multiple comparison test was used to compare the no-oligo control versus PNA molecule transfected cells. The PNA-peptide hybrid ASOs all showed a statistically significant reduction in the proportion of IR-B.

Discussion

The data presented in the gel (figure 3.2.3.B) and graphically in figure 3.2.3.C represents triplicate experiments using PNA-peptide hybrid ASO targeting the IR exon 11. The PNA ASO only, INSR PNA ex1, targeting the 5′ end of exon 11 (nucleotides 4-18) had no effect on IR exon 11 splicing (compare lane 2 and lane 4) at a concentration of 1 μM. The arginine/serine (RS) tail (INSR PNA tail only) transfected into HepG2 alone had no effect on exon 11 splicing (lane 5). The positive control, INSR ex2, continued to cause exon 11 skipping (lane 3). ASO INSR PNA ex1 5′ tail unexpectedly resulted in exon 11 skipping (lane 6) which was significantly increased in comparison to the no oligo control (lane 2) at 1 μM concentration. Figure 3.2.3.D shows a polyacrylamide gel from transfection with increasing concentrations of INSR PNA ex1 3′ tail. Increasing doses of INSR PNA ex 1 3′ tail from 250 nM (lane 3) to 1 μM (lane 5) also resulted in increased IR exon 11 skipping with a dose related reduction in the proportion of IR-B (figure 3.2.3.E) from to 13% to 6.6%.

Arginine/serine tails placed at either end of INSR PNA ex1 appeared to cause IR exon 11 skipping (compare 5′ tail - figure 3.2.3.B lane 7 and 3′ tail - 3.2.3.D lane 5). At equivalent concentrations, molecules with 3′ tails appeared to cause more exon skipping that those with 5′ tails (unpaired t test p= 0.0142).

2′-O-methyl RNA ASOs were more potent i.e. more efficacious at a lower concentration, compared with the PNA-peptide hybrid ASOs. For an equivalent effect on splicing, the concentration of INSR PNA ex1 3′ tail was 1 μM in comparison to 250 nM for the 2′-O-methyl RNA ASOs INSR ex2 (figure 3.2.3.D compare lanes 5 and 6).

Given that the tail sequence should provide a stimulus to enhancer splicing factors the PNA-peptide hybrid ASOs would be expected to enhance the IR exon 11

inclusion. On the basis of this evidence here modified ASOs with associated enhancer tails do not facilitate exon inclusion in the context of the IR exon 11. Possible explanations for these findings are discussed further in Chapter 4.

3.2.4 PNA-peptide hybrid targeting the 3′ end of the IR exon 11.

Background

The PNA-peptide hybrid ASOs, INSR PNA ex1 3′ or 5′ tail unexpectedly enhanced skipping of the IR exon 11 (Chapter 3.2.3). Despite PNA ASO targeting a neutral region of exon 11 (5′ end) and the theoretically splicing enhancer characteristics of an RS domain tail. To rule out the possibility that these effects were positional a further set of PNA-peptide hybrid ASOs were assessed. These targeted the 3′ end of exon 11. Though INSR ex2 caused exon 11 skipping and this further PNA-peptide hybrid ASO was to target a similar region, results from Chapter 3.2.2 demonstrate that the PNA INSR ex2 (18mer) had no effect on splicing.

The prediction was that the PNA-peptide hybrid ASO targeting the IR exon 11 with an enhancer tail sequence would increase its splicing. PNA-peptide hybrid molecules targeted the 3′ end of IR exon 11 with either a 3′ or a 5′ RS domain tails.

Aim

To assess the ability of PNA-peptide hybrid ASOs to enhance exon 11 inclusion by targeting its 3′ end.

Methods

The PNA-peptide hybrid targeting the 3′ end of exon 11 were manufactured by Bio-Synthesis (Texas, U.S.A). Essentially the PNA was a 15mer ASO targeting nucleotides 19 - 33 of the IR exon 11. The last 3 nucleotides of exon 11 were not targeted on the basis that this may hinder the 5′ splice site recognition by the spliceosome. The arginine/serine tail (10 repeats) was attached to the PNA ASO via a glycine-glycine linker. Two PNA-peptide hybrid ASOs were constructed with

the RS domain tail at either the 3′ end of the molecule or the 5′ end (figure 3.2.4.A) designated INSR PNA ex2 3′ tail and INSR PNA ex2 5′ tail respectively. Additional compounds where constructed to serve as controls and included the PNA ASO alone designated INSR PNA ex2 and the RS domain tail with the glycine-glycine linker alone designated PNA tail only. INSR PNA ex2 was a 15mer and is differentiated from the 18mer - INSR PNA ex2 (18mer) used in Chapter 3.2.2.

PNA-peptide hybrid ASOs were then transfected into HepG2 cells at varying concentrations for 48 hours as per the previously detailed methodology (Chapter 2 - materials and methods). The 2′-O-methyl RNA ASO, INSR ex2 was transfected as a positive control. RNA was subsequently harvested and subjected to RT-PCR. PCR primers allowed identification of the 2 isoforms of the IR and GAPDH as a control for RNA quantity and quality. Products were run on a polyacrylamide gel. Each experiment was replicated in triplicate.

Table 3.2.4.A provides replicate data of transfections involving ASO INSR ex2 from multiple experiments in previous chapters.

Results

Figure 3.2.4.A

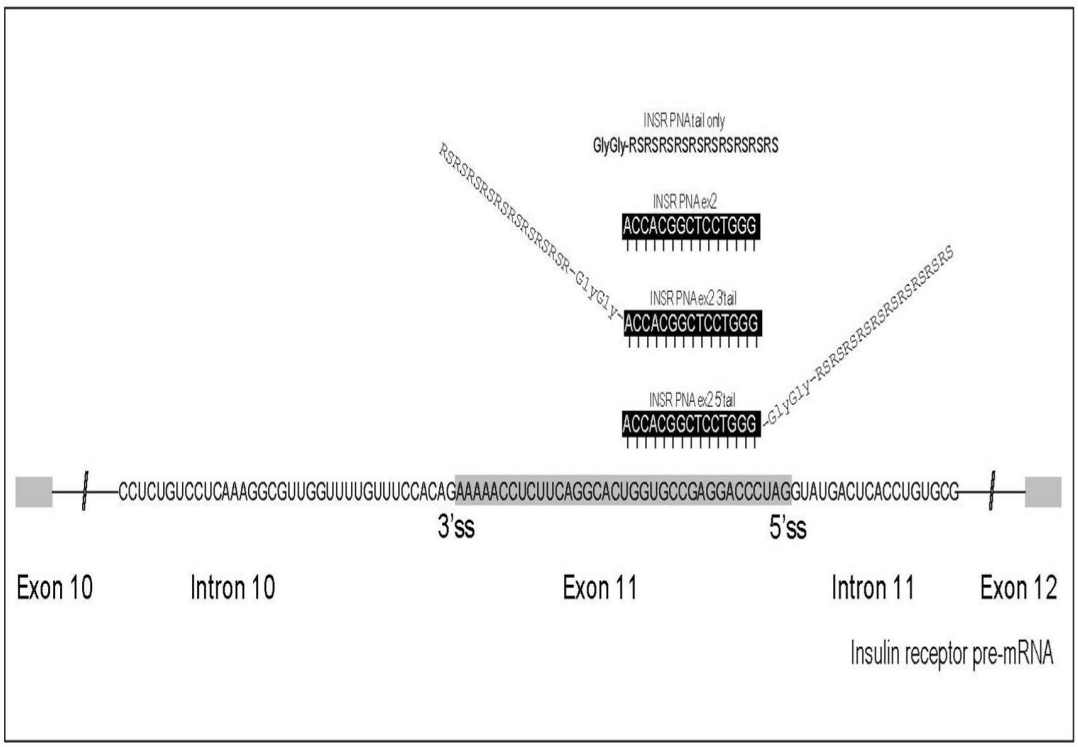


Figure 3.2.4.A PNA-peptide hybrid ASOs targeting the 3′ end of the IR exon 11. Schematic diagram shows a portion of the IR pre-mRNA from exon 10 to exon 12 (not to scale). Exon 11 and surrounding intronic sequences are shown. Exons are delineated by the grey shading and black characters, the rest of the sequence being intronic. PNA-peptide hybrid ASOs are illustrated above the IR pre-mRNA with black shadowing and white characters. The PNA backbone was 15 nucleotides long binding the 3′ end of exon 11 beyond the 3 nucleotides next to the intron/exon border. A tail with 10 arginine/serine (RS) repeats attached by a glycine-glycine link was attached either to the 3′ end of the PNA, designated INSR PNA ex2 3′ tail or the 5′ end of the PNA designated INSR PNA ex2 5′ tail. Both the arginine/serine tail alone (INSR PNA tail only) and the PNA alone (INSR PNA ex2) were transfected separately as controls. The ASO PNA ex2 (15mer) used in these hybrid ASOs was different to PNA INSR ex2 (18mer). Cells were transfected with ASOs at varying concentrations. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.

Figure 3.2.4.B

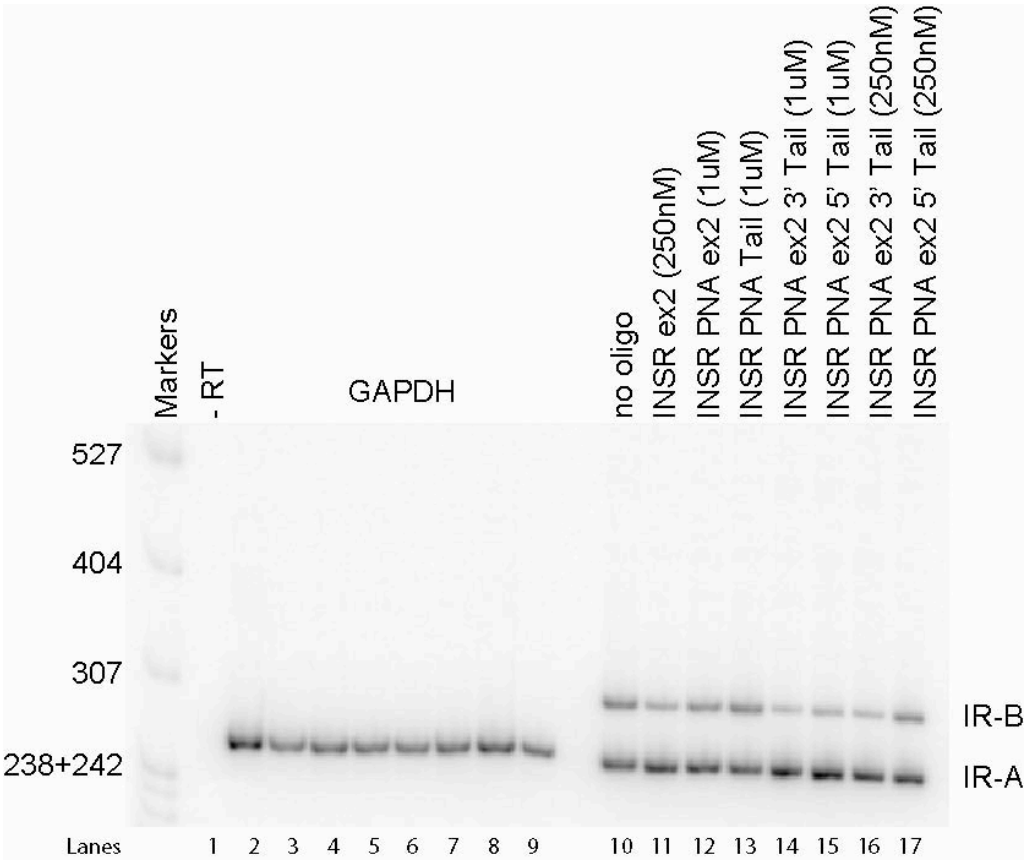


Figure 3.2.4.B PNA-peptide hybrid ASOs (3′ and 5′ tails) targeting the 3′ end of IR exon 11. Polyacrylamide gel showing IR isoforms from cells transfected with INSR PNA ex2 with RS domain tails and associated controls (no oligo, INSR PNA ex2 only, PNA tail only). ASO INSR ex2 was used as a positive control. IR-B (exon 11 inclusion) and IR-A (exon 11 exclusion) are shown. GAPDH was run as a control for RNA quality and quantity.

Figure 3.2.4.C

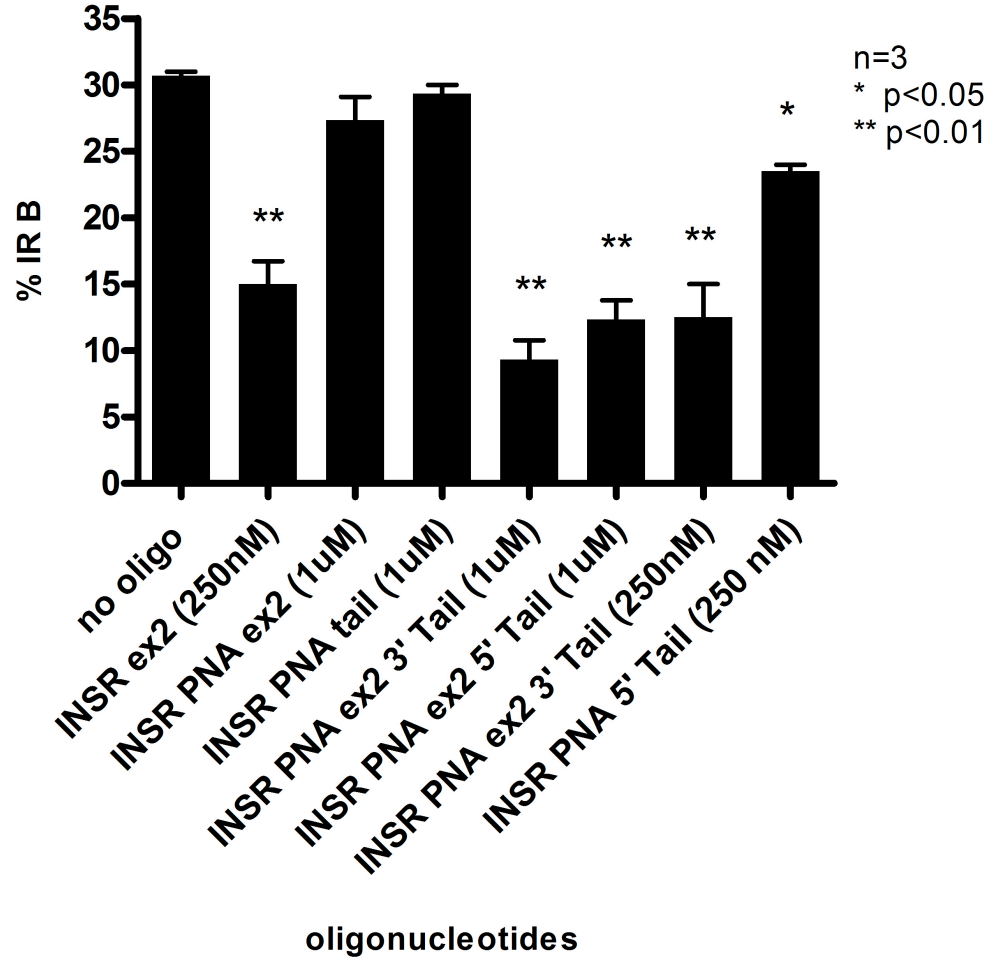


Figure 3.2.4.C PNA-peptide hybrid ASOs (3′ and 5′ tail) targeting the 3′ end of IR exon 11. The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection at varying concentrations (0-1μM). Quantitative data from three independent replications of the experiment (error bars show the S.E.M.). Overall one-way repeated measures ANOVA of the dataset showed $p<0.0001$, indicating statistically significant differences between groups. Dunnett's multiple comparison test was used to compare the no-oligo control versus PNA molecule transfected cells. The PNA-peptide ASOs all showed a statistically significant reduction in the proportion of IR-B.

Figure 3.2. 4.D

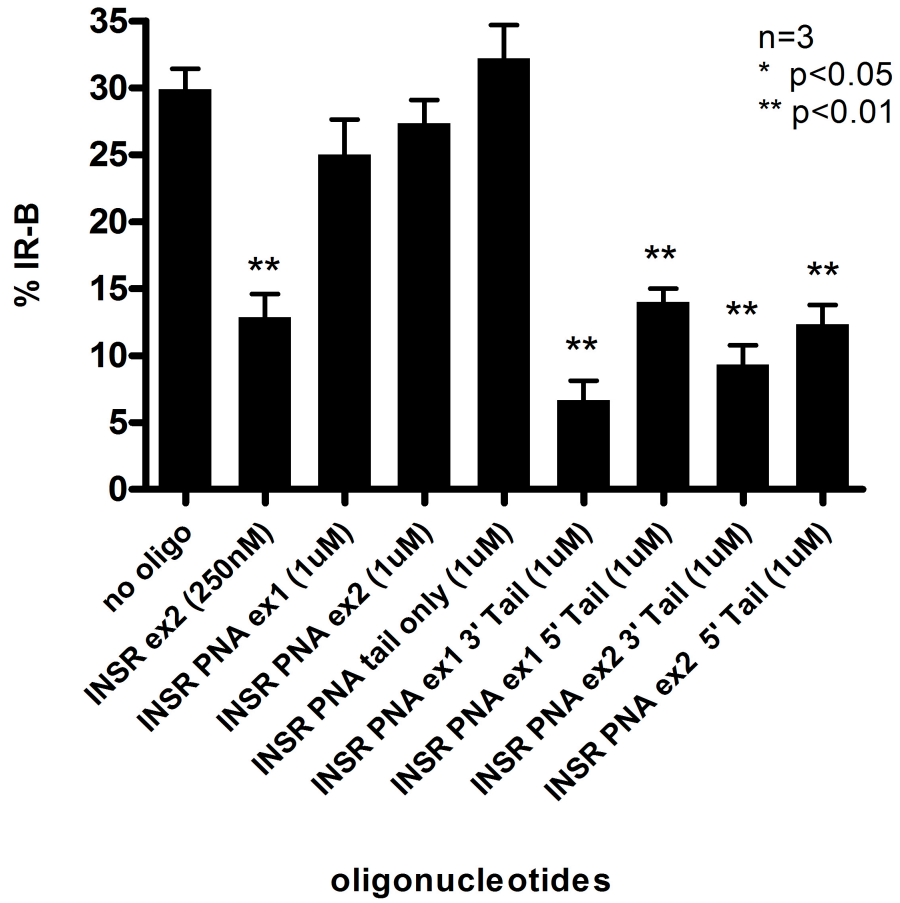


Figure 3.2.4.D PNA-peptide hybrid ASOs targeting the IR exon 11 - summary. The graph shows percentage IR-B with respect to total IR (IR-A + IR-B) following transfection at varying concentrations (0-1μM). Summary data for all PNA-peptide hybrid ASOs (error bars show the S.E.M.). Overall one-way repeated measures ANOVA of the dataset showed $p<0.0001$, indicating statistically significant differences between groups. Dunnett's multiple comparison test was used to compare the no-oligo control versus PNA molecule transfected cells. The PNA-peptide molecules all showed a statistically significant reduction in the proportion of IR-B as did the positive control ASO INSR ex2.

Table 3.2.4.A

IR-B/(IR-A+IR-B) %			
No oligo	INSR ex2	No oligo	INSR ex2
23	13	33	9
19	6	34	9
33	10	34	9
25	12	30	6.4
24	10	27	6
44	11	28	11
35	13	26	6
33	6	26	7
29	10	41	20
28	11	29	14
26	6	31	15
26	7	31	12
30	9	30	18

Table 3.2.4.A INSR ex2 ASO - Reproducibility. Data shown represents percentage IR-B as a proportion of total IR (IR-A+IR-B) from multiple replicates of transfection of ASO INSR ex2 into HepG2 cells. Results are given in two sets of columns with no oligo control transfection given in each case. The shaded cells represent data from transfection using a separate batch of INSR ex2. The no oligo transfection gave an IR-B/(IR-A+IR-B) of 29.8% (SD±5.35). The coefficient of variation is 18% for the replicates. IR-B/(IR-A+IR-B) for ASO INSR ex2 is 10.3% (SD±3.73) with a coefficient of variation of 36%. This improves to 27% when the data from the second batch of ASO INSR ex2 (shaded) is removed.

Discussion

Figure 3.2.4.B provides a representative gel from experiments involving the transfection of PNA-peptide hybrid ASOs targeting the 3′ end of the IR exon 11. Figure 3.2.4.C summarises graphically, the result of experiments carried out in triplicate.

HepG2 cells which were transfected with INSR PNA ex2 only (figure 3.2.4.B lane 12), at a concentration of 1 μM, revealed no change in the proportion of IR-B. Transfection with the INSR PNA tail only (lane 13), did not cause any alteration in the proportion of exon 11 splicing. These results are consistent with those noted in the previous Chapter 3.2.3.

PNA-peptide hybrid ASOs targeting the 3′ end of exon 11 including INSR PNA ex2 3′ tail (lanes 14 and 16) and INSR PNA ex2 5′ tail (lanes 15 and 17) failed to increase exon 11 inclusion and in fact caused exon 11 skipping. This effect increased with increasing ASO concentration from 250 nM to 1 μM (figure 3.2.4.C). This replicates data from Chapter 3.2.3 relating to the effects of PNA-peptide hybrids targeting the 5′ half of exon 11.

As with the PNA-peptide hybrids targeting the 5′ end of exon 11, the hybrid ASOs targeting the 3′ end of the exon appeared to be less potent than the 2′-O-methyl RNA ASOs at causing exon skipping. Equivalent effects between the compounds were seen when the PNA-peptide hybrids were 4 fold the concentration (1 μM) of the 2′-O-methyl RNA ASOs INSR ex2 (250 nM).

At a 1 μM concentration, INSR PNA ex2 3′ tail appeared to cause a greater reduction in the proportion of IR-B (9.3%) compared to INSR PNA ex2 5′ tail (12.3%) though this difference was not statistically significant (unpaired t test p= 0.22).

Figure 3.2.4.D graphically summarises the effects of all PNA-peptide hybrid ASOs tested in Chapter 3.2.3 and 3.2.4. This graph emphasises the finding that all PNA-peptide hybrid ASOs cause paradoxical IR exon 11 skipping. There also appears to be an inherent strength in a 3′ tail causing more exon skipping in comparison to a 5′ tail in PNA-peptide hybrids ASOs. This trend was noted in both sets of PNA-peptide ASOs tested but was significantly different in the INSR PNA ex1 3′ and 5′ tail transfected at a 1 μM concentration.

Figure 3.2.4.C shows the fall in the proportion of IR-B to 15% due to the effect of the ASO INSR ex2 at a concentration of 250nM. Table 3.2.4.A provides data from multiple replicates of INSR ex2 transfection in HepG2 cells taken from previous experiments. This demonstrates a mean reduction in IR-B to 10% overall. Two separate batches of INSR ex2 ASO were used in these sets of experiments. The latter batch of ASO INSR ex2 appeared to be less effective at causing exon skipping (shaded columns including data from this chapter).

Apparent differences in the effect of two batches of the same ASO (in this case the ASO INSR ex2) may be ascribed to either variation in experimental conditions or the ASO used. Assuming identical experimental conditions, then the ‘purity’ of the ASO in terms of the presence of full length transcript only, sequence fidelity and the presence of contaminates (by-products of the synthesis process) may account for variable effects seen. Failure sequences (oligo sequences aborted during synthesis) or those with an incorrect sequence may not exert the same effects in cell culture as the full length correctly sequenced transcript. Contaminants used in the synthesis and purification process may be toxic in cell culture.

The ASOs used were commercially acquired (Dharmacon). Standard purification methods include desalting to remove the by-products of the synthesis process

either using gel infiltration or ethanol precipitation. Short of re-sequencing the ASO to check sequence fidelity, MALDI-TOF mass spectrometry is used as a surrogate marker to confirm that the full length transcript has been generated. With standard production methods wrongly sequenced ASOs are unlikely. An 85% purity level is quoted for full length ASO production. Purification via HPLC (high pressure liquid chromatography) or PAGE (polyacrylamide gel electrophoresis) can remove failure sequences and can increase purity up to 99% but was not used in the case.

Chapter 3.3

2′-O-methyl RNA antisense oligonucleotides targeting multiple exonic silencer sequences within Apolipoprotein B exon 26

3.3 2′-O-methyl RNA ASOs targeting multiple exonic splicing silencer within Apolipoprotein B exon 26.

Background

The average length of an exon has been estimated at 137 nucleotides (Hawkins 1988). Therefore, at 7,572 nucleotides, the constitutively spliced APOB exon 26 is unusually long. This is in stark contrast to the short, alternatively spliced IR exon 11 at 36 nucleotides which had been the subject of study in Chapter 3.1 and 3.2. There are two predominant splice site sequences to which the majority sites conform in canonical splicing. At the 5' splice site it is the sequence GU and AG at the 3' splice site. However there are exceptions. A non-consensus sequence at 5' splice site is the dinucleotide GC (Wu et al. 1999) though these splice sites appear to match well in other positions. U12 introns have different conserved sequences (Tarn et al. 1997). There are also other non-consensus cryptic splice sites which are recognised when the native splice site is inactivated without the dinucleotides AG or GU (Roca et al. 2003). Consensus sequences defining the 5′ and 3′ splice sites of these exons are highly degenerate (Shapiro et al. 1987). As a result, matches to these consensus sequences are highly prevalent throughout the genome (Fairbrother et al. 2000) and outnumber genuine splice sites by an order of magnitude e.g. the human HPRT gene (Sun et al. 2000). These 'pseudo splice sites' are not normally employed in splicing but it remains unclear how the spliceosome is able to differentiate these numerous sites from few genuine splice sites present. Due to its length, APOB exon 26 contains multiple pseudo splice sites which are not spliced.

However further study of these pseudo splice sites using the position score matrix of Shapiro and Senapathy (Shapiro et al. 1987) raises a paradox. Position score

matrices out of 100 provide an estimate of the likelihood of a sequence being identified as an exon boundary and utilised in splicing. The 3′ splice site flanking exon 26 scored 83.2 and the 5′ splice site scored 86.5 (combined score 169.7). Using the exon 26 flanking scores of 83.2 and 86.5 as thresholds, 32 3′ splice sites and two 5′ splice sites within exon 26 were identified as closer matches to the consensus with higher scores than the genuine splice sites flanking exon 26. On theoretical grounds, therefore, each of these pseudo splice sites could be used in preference to the genuine splice sites. Two possible 3′ shortened alternatives combining the 5′ pseudo splice sites with the 3′ splice site flanking exon 27, and 32 possible 5′ shortened alternatives combining the 3′ pseudo splice sites with the 5′ splice site flanking exon 25.

Further to this, a search was also made for plausible pseudoexons within the length of exon 26 using this subset of high-scoring pseudo splice sites. These were defined as sequences bounded by 3′ and 5′ pseudo splice sites in the right orientation, with lengths of 39-374 base pairs, these being the range of lengths for the other APOB internal exons. Using these criteria at least 3 alternative pseudoexons could be identified that contain pseudo splice sites that are a better fit to the consensus than the genuine splice sites, and which on these bioinformatic criteria are more plausible exons than exon 26 itself.

This analysis raises the issue of why certain pseudo splice sites are not spliced in preference to the apparently weaker constitutive sites. Miriami E et al. have shown that latent 5’ splice site in introns have a higher frequency of associated stop codons in their vicinity than introns without these sites suggesting an inbuilt mechanism to suppress their use (Miriami et al. 2004). Another hypothesis is that exonic splicing silencer (ESS) sites may suppress the use of these potential pseudo sites. Previous work undertaken using the in vitro heterologous splicing

reporter assays DNA ligase III and ß-globin splicing reporter HßΔ6 has attempted to address this question. Candidate sequences, initially 400 nucleotides in length where derived from exon 26 and cloned into the reporters replacing the native ESS. These constructs were transcribed and tested for splicing in vitro. Suppression of C-ß or HßΔ6 splicing indicates the presence of ESS activity within the candidate sequence. Various sequences of 25 nucleotides, derived from the 400 nucleotide sequences demonstrating silencing activity, were then cloned into the splicing reporter to investigate these regions more closely. This work identified multiple sequences along the length of exon 26 with silencer activity suggesting the presence of multiple ESSs (unpublished data). It remains unclear whether these multiple sites work independently or in tandem and indeed how physiologically relevant they might be in the in vivo cellular environment.

To investigate the role of ESS elements in in vivo cell culture, we turned to 2′-O-methyl RNA ASOs targeting some of these regulatory elements identified in previous experiments. ASOs provided an opportunity to test the effects of blocking specific sites in isolation and in tandem to assess the effects in splicing.

The prediction was that ESS sites in APOB exon 26 allow the exon to be constitutively spliced by suppressing pseudo splice sites within it. Blocking these ESS sites in APOB exon 26 using 2′-O-methyl RNA ASOs may alter its splicing choices preferentially favouring a stronger pseudo splice site within it.

Aim

To assess the effect of targeting ESS elements separately and in tandem on the splicing of APOB exon 26 using 2′-O-methyl RNA ASOs.

Methods

We chose to use ASOs in an in vivo model with the native APOB transcript to provide proof of principle in the most physiological manner possible. 2′-O-methyl RNA ASOs were designed to sites predicted to behave as ESSs in APOB exon 26 from previous studies. ASOs were designed as 25mers in line with the length of sequences assessed in the in vitro assays. ESS elements were targeted from 2 regions of exon 26. Firstly within the first 200 nucleotides of 3′ splice site of APOB exon 26, designated A1 and a region approximately 2300 nucleotides from the 3′ splice site in a region designated A6. Regions A1 and A6 were chosen to provide a contrasting analysis of two regions with both low and high silencing activity respectively. The A1 region has fewer silencing sequences that A6 but within this region (A1) there are powerful silencer regions between nucleotides 65-89. The specific targeted regions identified as ESS were at nucleotides 66-89, 97-121, 129-153, 2319-2343 designated APOB ex1, APOB ex2, APOB ex3 and APOB ex4 (figures 3.3.A; 3.3.B and 3.3.C). Within regions A1 and A6, 2 neutral sequences were targeted with ASOs as negative controls. The ASO APOB ex5 targeted the neutral A1 region 193-217 while APOB ex6 targeted the neutral A6 region 2191-2215.

These ASOs were transfected into HepG2 cells, which are a well-characterised cell line that synthesises and secretes Apolipoprotein B. The transfection agent lipofectamine 2000 (Gibco/Invitrogen, Paisley, UK) was used. ASOs were transfected individually or in combinations (at 125 nM concentration each). In order to equalise the total concentration of ASO per reaction (500 nM) a neutral ASO, in this case INSR ex2scram, was used to make up the differences. INSR ex2scram was also transfected alone to provide a further negative control ASO which did not match any sequences within the APOB pre-mRNA. ASOs were incubated with

HepG2 cells for 48 hours at 37^{o}C/5% CO_2. RNA was subsequently harvested and subjected to RT-PCR to assess the splicing of APOB exon 26. PCR was undertaken with the EXL DNA polymerase (Stratagene, CA, USA) which was particularly optimised for replication of longer targets (greater than 20 kb) and with higher fidelity than Taq polymerases. PCR primers were designed to allow the identification of the full length exon 26 transcript with forward and reverse primers being located in APOB exon 25 and APOB exon 27 respectively (primer sequences given in Chapter 6.2 - appendix). Products were run on a 6% polyacrylamide gel. The PCR incorporated the radioactive nucleotide [α-^{32}P] dCTP (Perkin-Elmer, Wellesley, USA) allowing relevant bands to be identified and quantified using an autoradiograph. Each experiment was replicated in triplicate. Bands which appeared to be new or increased in the presence of ASOs were identified and removed from the gel. The remaining bands were not sequenced. Bands were then purified using a PCR purification kit (Qiagen, Valencia, CA) and a further PCR undertaken without radioactive nucleotides in order to send products for sequencing. Products were sent for sequencing via the Genome centre, William Harvey Research Institute. An ABI 3700 automated DNA/RNA Sequencer (Applied Biosystems) was used, employing the ABI Prism Big Dye sequencing Kit. Data Collection software 2.0 was used to collect data and results were analysed with the BioEdit sequence analysis software version 7. Expression of novel bands was reported as the relative expression of a given band in comparison to its expression in the no oligo control lane.

Figure 3.3.A

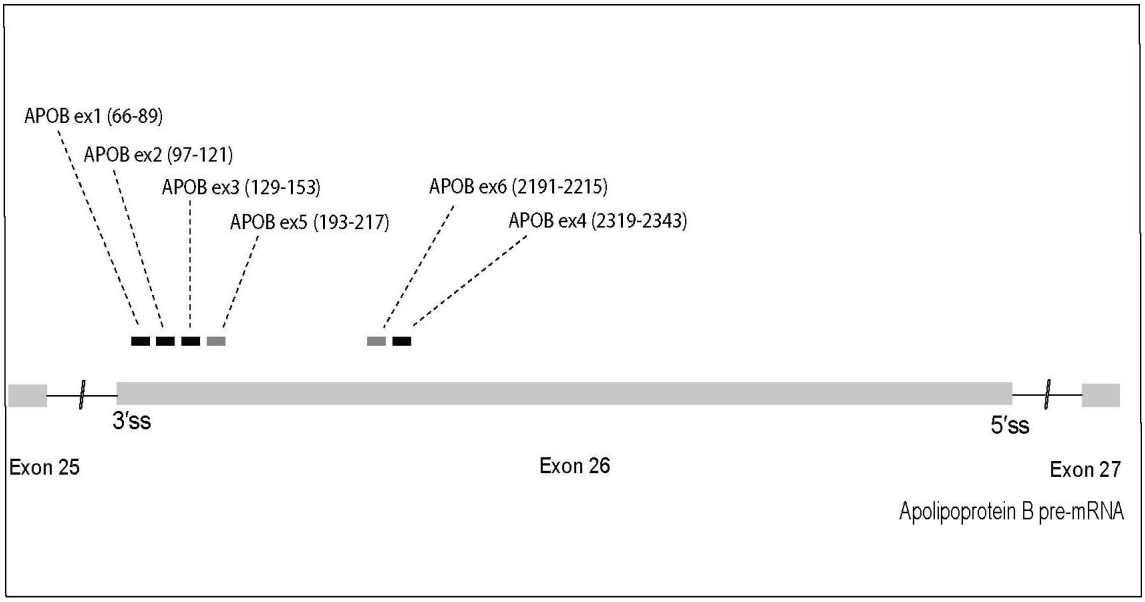


Figure 3.3.A 2′-O-methyl RNA ASOs targeting ESSs in APOB exon 26. Schematic diagram shows APOB pre-mRNA including exon 25, 26 and 27 (not to scale). 2′-O-methyl RNA ASOs (25mers) were designed to bind putative exonic splicing silencer (ESS) sites in exon 26 as shown. Four ASOs, designated APOB ex1; APOB ex2; APOB ex3 and ApoB ex4 blocked ESS sites (shown as black blocks). APOB ex5 and APOB ex6 blocked neutral sites as controls (shown as small grey blocks).

Figure 3.3.B

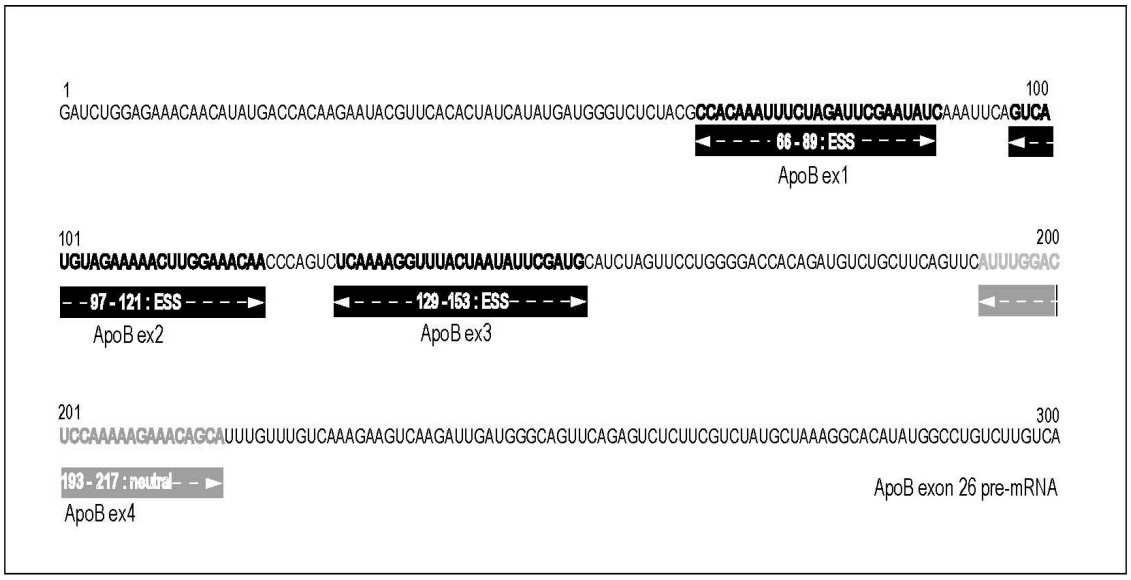


Figure 3.3.B 2′-O-methyl RNA ASOs targeting ESSs in region A1 of APOB exon 26. Schematic diagram shows the 5′ end of the ApoB exon 26 pre-mRNA (figures at each end of the nucleotide sequence indicate position from 5′ end of APOB exon 26). Sequences identified as ESS are delineated by bold type and neutral sequences by light grey type. 2′-O-methyl RNA ASOs, 25 nucleotides long, were targeted to 3 predicted ESS sequences at the 5′ end of ApoB exon 26, designated APOB ex1, APOB ex2, APOB ex3 (black boxes, white characters). A further 2′-O-methyl RNA ASO, 25 nucleotides long, was designed to target the neutral sequence, designated APOB ex4 (grey box, white characters). HepG2 cells were transfected with ASOs in varying combinations and at varying concentrations. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.

Figure 3.3.C

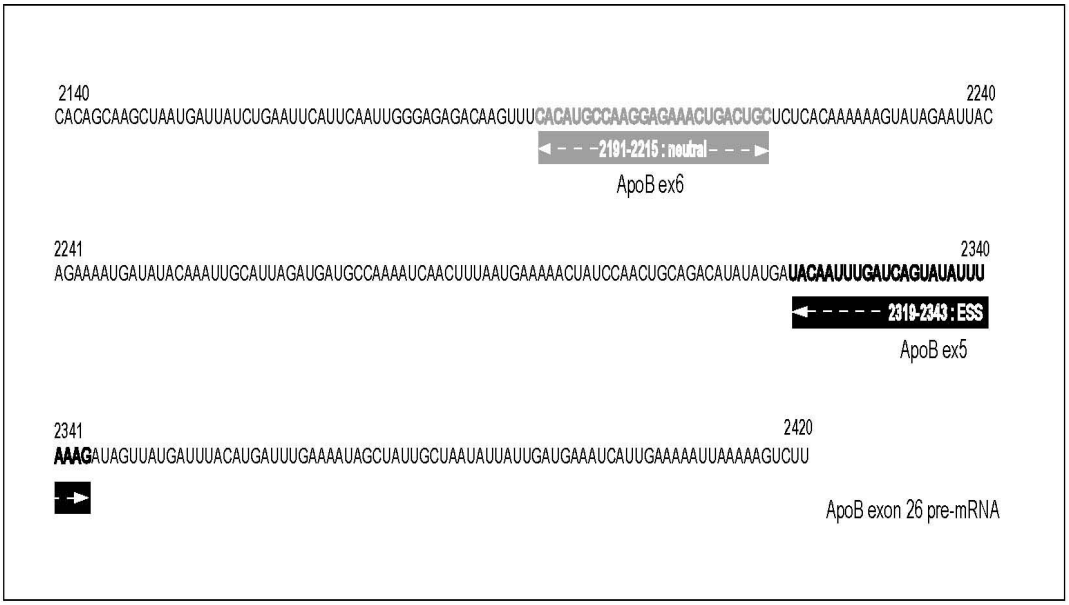


Figure 3.3.C 2′-O-methyl RNA ASOs targeting ESSs in region A6 of APOB exon 26. Schematic diagram shows a region in the first third of the ApoB exon 26 pre-mRNA (figures at each end of the nucleotide sequence indicate position from 5′ end of ApoB exon 26). Sequences identified as ESS are delineated by bold type and neutral sequences by light grey type. A 2′-O-methyl RNA ASO, 25 nucleotides long, was targeted to an ESS sequences in the A6 region of APOB exon 26, designated APOB ex5 (black boxes, white characters). A further 2′-O-methyl RNA ASO, 25 nucleotides long, was designed to target a neutral sequence, designated APOB ex6 (grey box, white characters). HepG2 cells were transfected with ASOs in varying combinations and at varying concentrations. RT-PCR was carried out on the total RNA extracted from these cells after 48 hours.

Figure 3.3.D

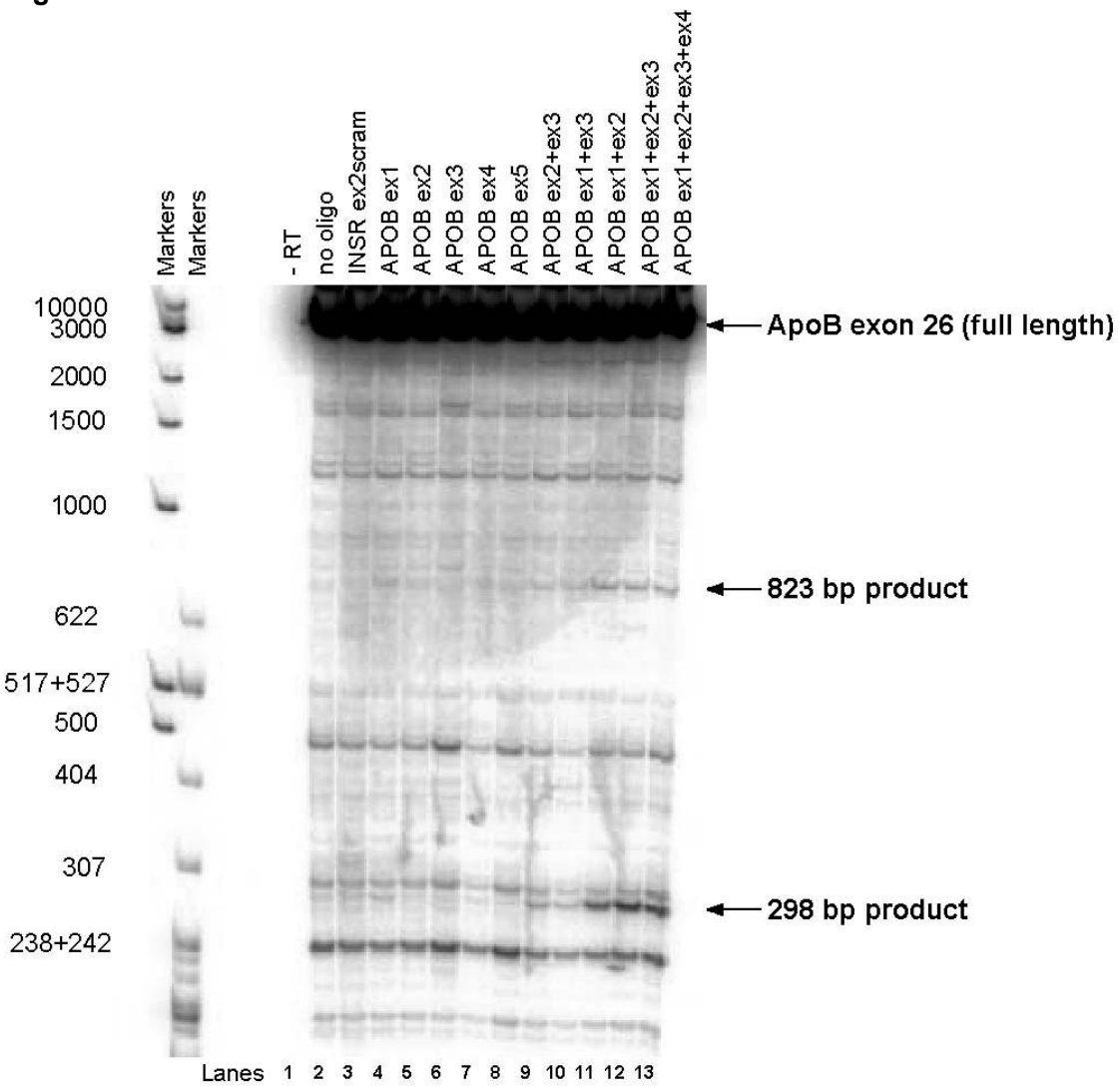


Figure 3.3.D 2′-O-methyl RNA ASOs targeting ESSs in APOB exon 26. Representative gel from HepG2 cells transfected with single and combination ASOs targeting ESS in APOB exon 26. Each lane represents a well of transfection. Fragment lengths in bp of the markers are indicated on the left side. The -RT control PCR (lane 1) was performed without reverse transcriptase. In lane 2 (no oligo) the cells were not transfected with ASOs. In lane 3, the cells were transfected with negative control ASO at 500 nM concentration. Lanes 4-13 were transfected with the indicated ASOs at 125 nM each; the total concentration of ASO was adjusted to 500 nM with additional negative control ASO. APOB ex1, ex2, ex3 and ex4 targeted putative ESS sequences while APOB ex5 was a control ASO targeting a neutral region. Cells were incubated for 48 hours and RT-PCR was carried out on the total RNA extracted from these cells with primers annealing to the adjacent exons 25 and 27. The positions of the bands corresponding to the APOB mRNA species with constitutive exon 26 inclusion (top), pseudo 3′ splice site activation (middle) and skipping of exon 26 (bottom) are indicated on the right side.

Figure 3.3.E

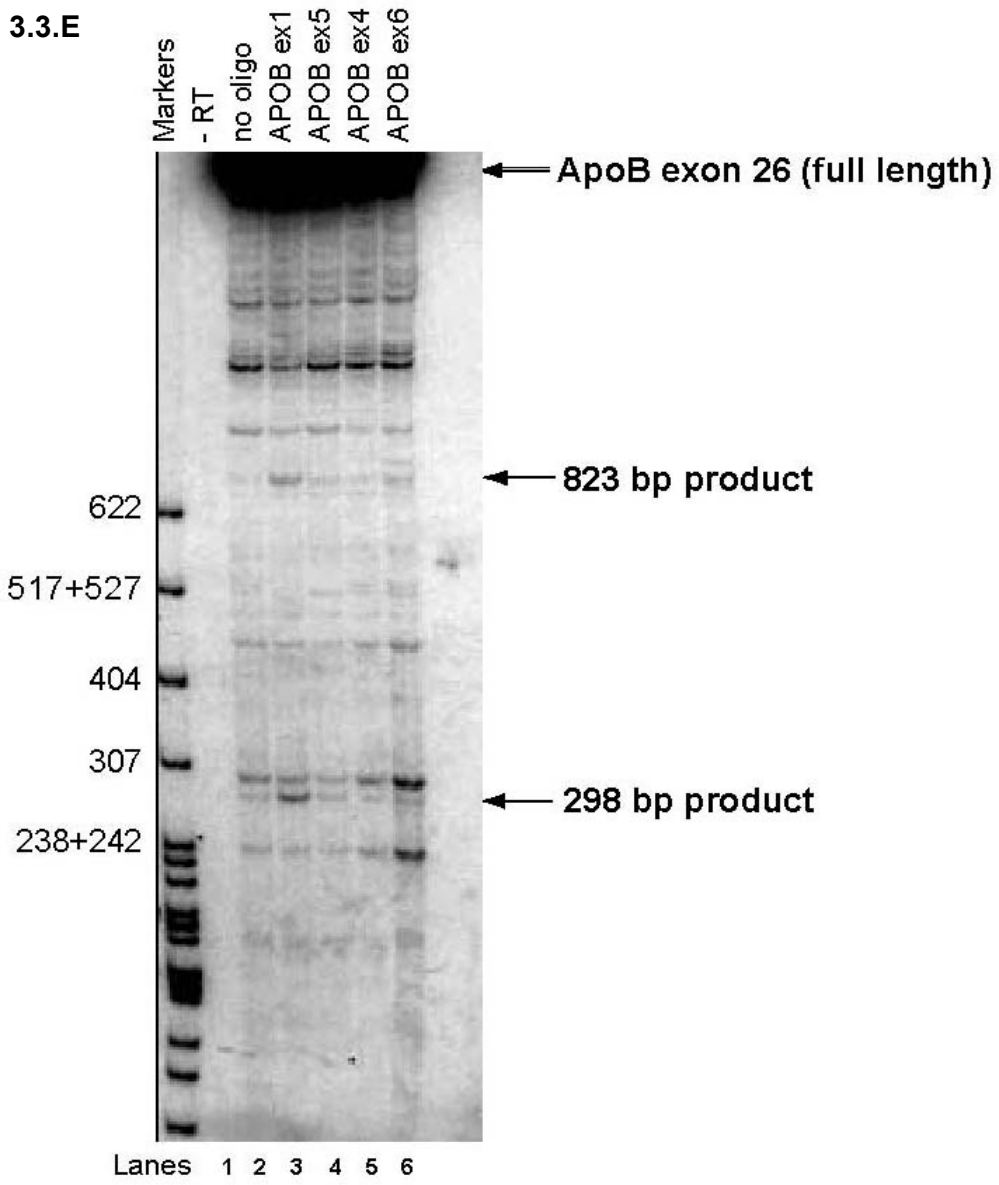


Figure 3.3.E 2′-O-methyl ASOs targeting ESSs in APOB exon 26. Representative gel from HepG2 cells transfected with single ASOs targeting ESS and neutral sequences in APOB exon 26. Each lane represents a well of transfection. Fragment lengths in base-pairs of the markers are indicated on the left side. The -RT control PCR (lane 1) was performed without reverse transcriptase. In lane 2 (no oligo) the cells were not transfected with ASOs. Lanes 3-6 were transfected with the indicated ASOs. This figure shows transfections with 2 ASOs targeting ESS sites and 2 ASOs targeting neutral sequences. APOB ex1 and ex4 targeted putative ESS sequences while APOB ex5 and APOB ex6 were control ASOs targeting neutral regions. This figure was principally included to show the effects of the second neutral sequence APOB ex6. Cells were incubated for 48 hours and RT-PCR was carried out on the total RNA extracted from these cells with primers annealing to the adjacent exons 25 and 27.

Figure 3.3.F

823 base product

GCCATCTCGAGAGTTCCAAGTCCCTACTTTTACCATTCCCAAGTTGTATCAACTGCAAGTGCCT
CTCCTGGGTGTTCTAGACCTCTCCACGAATGTCTACAGCAACTTGTACAACTGGTCCGCCTCCT
ACAGTGGTGGCAACACCAGCACAGACCATTTCAGCCTTCGGGCTCGTTACCACATGAAGGCTG
ACTCTGTGGTTGACCTGCTTTCCTACAATGTGCAAGATCTTCAGGTTCCATCGTGCAAACTTGAC
TTCAGAGAAATACAAATCTATAAGAAGCTGAGAACTTCATCATTTGCCCTCAACCTACCAACACT
CCCCGAGGTAAAATTCCCTGAAGTTGATGTGTTAACAAAATATTCTCAACCAGAAGACTCCTTGA
TTCCCTTTTTTGAGATAACCGTGCCTGAATCTCAGTTAACTGTGTCCCAGTTCACGCTTCCAAAA
AGTGTTTCAGATGGCATTGCTGCTTTGGATCTAAATGCAGTAGCCAACAAGATCGCAGACTTTGA
GTTGCCCACCATCATCGTGCCTGAGCAGACCATTGAGATTCCCTCCATTAAGTTCTCTGTACCT
GCTGGAATTGTCATTCCTTCCTTTCAAGCACTGACTGCACGCTTTGAGGTAGACTCTCCCGTGTA
TAATGCCACTTGGAGTGCCAGTTTGAAAAACAAAGCAGATTATGTTGAAACAGTCCTGGATTCCA
CATGCAGCTCAACCGTACAGTTCCTAGAATATGAACTAAATGTTTTGGGAACACACAAAATCGAA
GATGGTACGTTAGCCTCTAAGACTAAAG**GAACATTTGCACACCGTGAC**

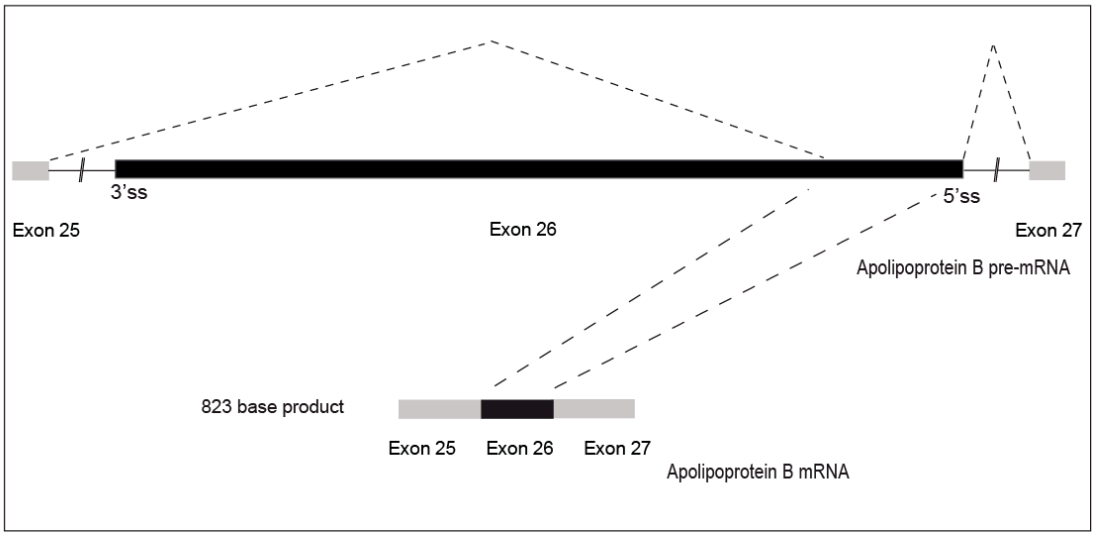


Figure 3.3.F 2′-O-methyl ASOs targeting ESSs in APOB exon 26 - 823 base alternatively spliced product. Sequence of the 298 base product. Black lettering at each end of the sequence signifies primer sequences. Central sequence black lettering - exon 26. Light grey sequence - exon 25. Dark grey sequence - exon 27. Alternative 3' splice site selected (see schematic lower figure).

Figure 3.3.G

298 base product

GCCATCTCGAGAGTTCCAAGTCCCTACTTTTACCATTCCCAAGTTGTATCAACT
GCAAGTGCCTCTCCTGGGTGTTCTAGACCTCTCCACGAATGTCTACAGCAACTT
GTACAACTGGTCCGCCTCCTACAGTGGTGGCAACACCAGCACAGACCATTTCA
GCCTTCGGGCTCGTTACCACATGAAGGCTGACTCTGTGGTTGACCTGCTTTCCT
ACAATGTGCAAGTTTTGGGAACACACAAAATCGAAGATGGTACGTTAGCCTCTA
AGACTAAAG**GAACATTTGCACACCGTGAC**

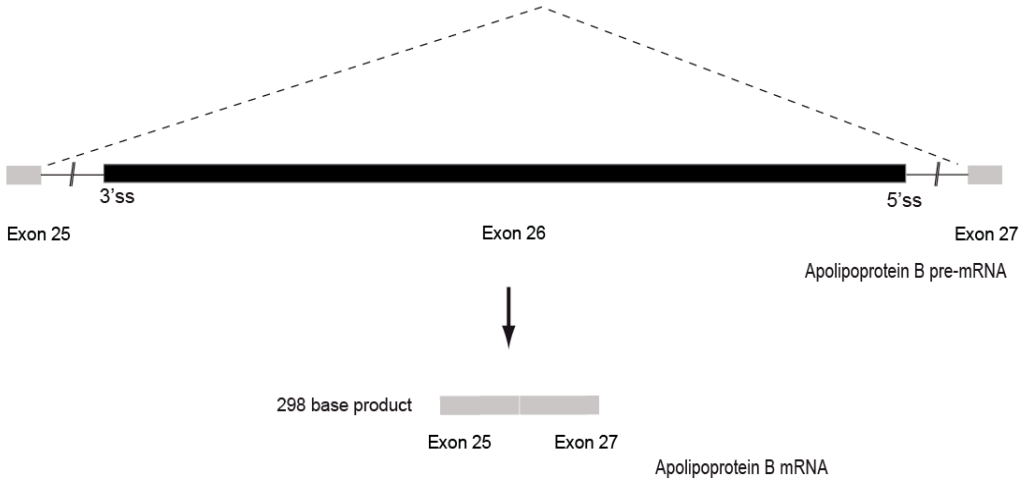


Figure 3.3.G 2′-O-methyl ASOs targeting ESSs in APOB exon 26 - 298 base alternatively spliced product. Sequence of the 298 base product. Black lettering signifies primer sequences. Light grey sequence - exon 25. Dark grey sequence - exon 27. Exon 26 entirely skipped (see schematic lower figure).

Figure 3.3.H

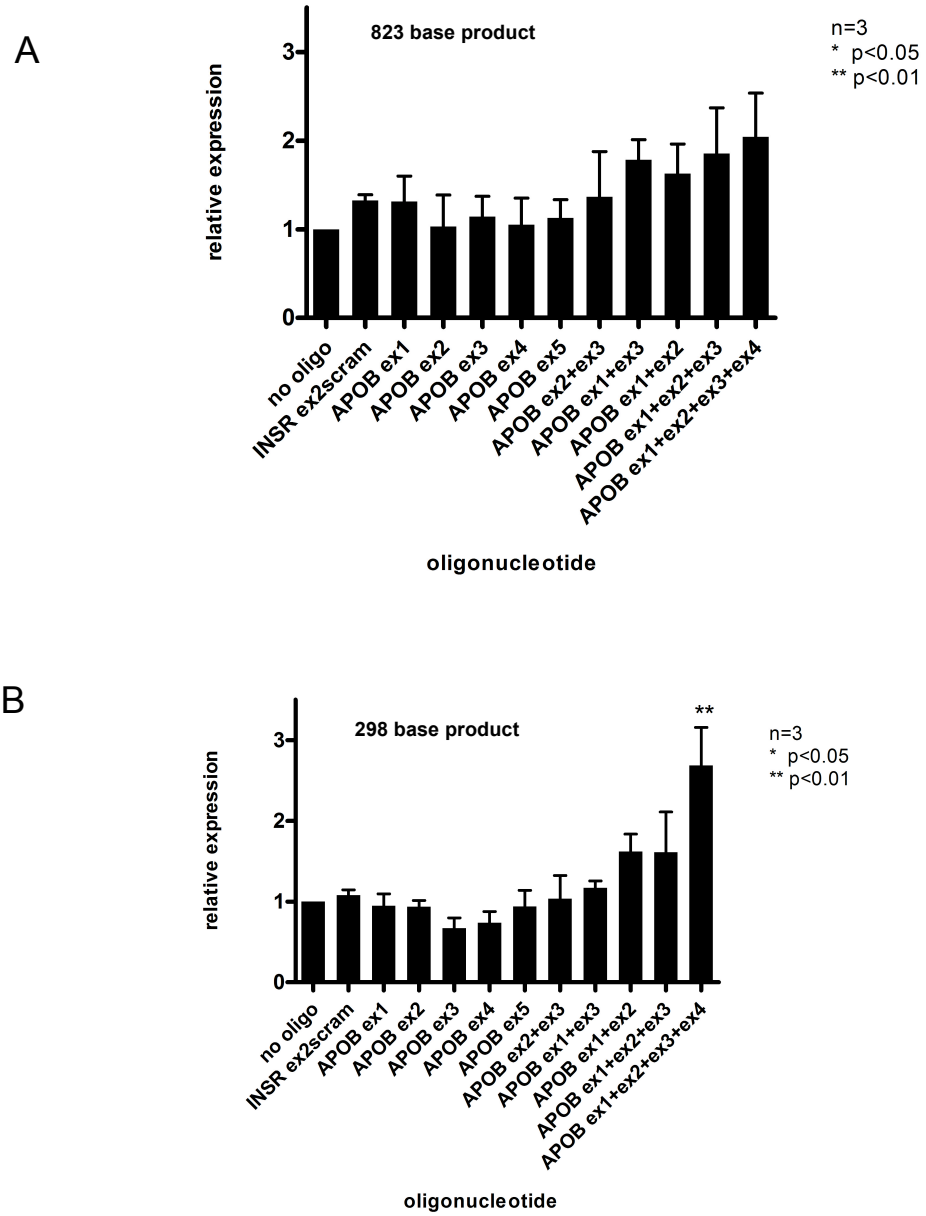


Figure 3.3.H 2′-O-methyl RNA ASOs targeting ESSs in APOB exon 26 - alternatively spliced products. The graphs show the relative expression of alternatively spliced products in proportion to their expression in the no oligo control lane. Each ASO was transfected at a concentration of 125nM. Quantitative data from three independent replications of the experiment (error bars show the S.E.M.). **A.** Relative expression of 823 base product. Overall one-way repeated measures ANOVA of the dataset showed p=0.3848 indicating no statistical difference between the groups. **B.** Relative expression of 298 base product. Overall one-way repeated measures ANOVA of the dataset showed p=0.0004 indicating statistically significant differences between groups. Addition of APOB ex1 + ex2 + ex3 and ex4 resulted in a statistical significant increase in the relative expression of the 298 base product.

Discussion

Figures 3.3.D and 3.3.E were polyacrylamide gels from representative experiments in which 2′-O-methyl RNA ASOs targeting ESS sequences in APOB exon 26 were transfected in cell culture individually or in combination with other ASOs. There were a large number of splice products for each reaction. The constitutive splicing of exon 26 at the native splice sites continues to generate the dominant isoform. As shown in both figures 3.3.D and 3.3.E, targeting ESSs individually did not cause significant alternative splicing (figure 3.3.D compare lanes 2 and 3 with lanes 4-7). Targeting the neutral sequence 193-217 with ASO APOB ex5 (figure 3.3.D - compare lanes 2 and 3 with lane 8) or the neutral sequence 2191-2215 with ASO APOB ex6 did not alter splicing (figure 3.3.E - compare lanes 2 and lane 6). Though not identical in terms of base composition, an unrelated scramble sequence (INSR ex2scram) was included. The scramble sequence confirms that the effects on splicing were not dependent merely on the addition of oligonucleotide alone. No specific APOB scram sequences were tested but would have provided a more robust control.

When mixtures of ASOs targeting ESSs were combined, increased alternative splicing was seen (figure 3.3.D - compare lanes 2 and 3 with lanes 9-13). Two alternative isoforms were noted. A 823 base pair (figure 3.3.F) and a 298 base pair product (figure 3.3.G) appeared to be increased when multiple ESSs were blocked. These two products were purified, subjected to a further PCR and then sequenced. The 823 base product was found to define an alternative 3′ splice site, 7047 bases downstream of the constitutive 3′ splice site (Shapiro and Senapathy score 90.4). The 298 base product was found to be as a result of complete exclusion of exon 26.

The amounts of the alternatively spliced products identified on transfection of various combinations of ASOs against APOB exon 26 were represented as a proportion of that spliced product when no oligonucleotide was transfected. These results are shown graphically in figure 3.3.H. Not all oligonucleotides resulted in more alternative splicing i.e. alternative spliced product > 100% (see figure 3.3.H B APOB ex3 and APOB ex4).

The products were not represented as a proportion of APOB for two reasons. Firstly, as the bands for total APOB were saturated the gel quantification was inaccurate. Secondly differentiating between the effects of combinations of ASOs would be difficult given the small amount of alternative product as a proportion of the total APOB.

Most of the truncated forms of APOB would be degraded by mechanisms such as nonsense mediated decay. However we do know that some truncated isoforms of APOB are stable and avoid degradation. The truncated form APOB48 is stably present in the intestine and undertakes a crucial role in chylomicron function. Heterozygote patients with hypobetalipoproteinemia produce the truncated APOB form APOB89 (Talmud et al. 1989). Transfected ASOs heve been shown to generated a truncated APOB transcript, APOB87, which was shown to result in a stably expressed protein (Khoo et al. 2007).

These results confirmed the prediction that in vivo blocking of selected ESS with ASOs trigger the selection of a strong 3′ pseudo splice site within exon 26 i.e. cryptic activation of the pseudo splice site. Constitutive splicing of exon 26 at the native splice sites continues to generate the dominant isoform and the alternative splicing effects are small. This study suggests that ESS elements have a function in suppressing pseudo splice site selection in vivo. These experiments also

demonstrated that ESS sites were functionally redundant in that that only by disabling multiple sites could a strong 3′ pseudo splice site be activated.

Chapter 4

Discussion

4.1 Summary of main biological findings.

Both the IR and APOB genes code proteins involved in central metabolic processes. The study of the regulation of these genes provides models of splicing in a short, alternatively spliced exon (IR exon 11) and a long constitutive exon (APOB exon 26). ASO technology provides a strategy to study splicing mechanisms in these diverse exons and the opportunity for future therapeutic intervention as understanding of the technology expands.

This current work provides complimentary, conflicting and new data to this area. The significant findings of this study add to the current body of knowledge in two main areas, namely; splicing in the IR exon 11 and APOB exon 26 and the use of ASO technology to alter splicing.

4.1.1 2′-O-methyl RNA ASOs targeting the IR exon 11.

Using 2′-O-methyl RNA ASOs, intronic and exonic sites identified as being important to the splicing of the IR exon 11 were targeted. Intronic sites included the 3′ splice site, the 5′ splice site, the branch point sequence and combinations of these. Exonic sites included the 5′ half (INSR ex1), 3′ half (INSR ex2), central region (INSR ex3) and the whole of exon11 (INSR ex1+2). This data suggested blocking intronic sites individually has little effect on splicing but that blocking combinations of intronic sites causes exon 11 skipping. ASOs blocking the 3′ end (INSR ex2) and the whole of the exon (INSR ex1+2) cause exon 11 skipping whereas those blocking the 5′ half (INSR ex1) and the central region (INSR ex3) have no effect. This technique therefore suggests an enhancer region at the 3′ end of exon 11. The remainder of the exon contains either no further modulating regions or a combination of opposing regions.

4.1.2 TOES ASOs targeting the IR exon 11.

TOES ASOs designed to enhance IR exon 11 inclusion had no effect on the splicing of this exon. The lack of effect may have been the result of:

1. Failure to localise or bind to the IR pre-mRNA transcript as a result of poor transfection efficiency. The positive control in the experiments suggested that transfection had occurred appropriately. Given the charged components within the TOES ASO, transfection across the polar cellular membrane would not be expected to be hindered.
2. Steric interference – TOES ASOs may sterically interfere with interaction between the spliceosome and IR exon 11.
3. Secondary structure effects – TOES ASOs bind IR exon 11 and alter pre-mRNA secondary structure in such a way as to inhibit splicing.
4. Incorrect recruitment of splicing factors by ESE tail.

Studies may be designed to clarify these issues.

1. Poor transfection efficiency/binding

    - Assess transfection efficacy using fluorescent - tagged ASOs

    - Change transfection conditions e.g. transfection agent etc.

    - Gel shift binding assay comparing TOES compound with a 2'O-methyl ASO shown to alter splicing. A gel shift binding assay utilising an in vitro splicing reaction including radiolabelled ASOs to compare and a pre-mRNA transcript including the IR exon 11 could be undertaken. The bound and unbound ASO could then be resolved by separating on a polyacrylamide gel and quantifying.

5. Steric interference – co-immunoprecipitation study assessing the interference of binding of spliceosomal elements by TOES ASOs
6. Secondary structure effects – Utilise plasmid which contain mutated IR exon 11 sequences designed to be resistant to secondary structure formation (chapter 4.5.1) whilst retaining the ability to bind ASOs. Assess whether TOES ASO affect splicing in the absence of secondary structure.
7. Incorrect recruitment of splicing factors by ESE tail – Redesign TOES ASO with different ESE tails.

4.1.3 PNA and PNA-peptide hybrid ASOs targeting the IR exon 11.

4.1.3.1 Equivalent PNA and 2′-O-methyl RNA ASOs do not have equivalent effect on splicing in the IR exon 11 splicing.

Experiments from Chapter 3.2 demonstrate that PNA ASOs alone, targeting the IR exon 11 had no effect on its splicing. PNA ASOs only demonstrated an effect on splicing as part of the compound PNA-peptide hybrid ASOs. The 2′-O-methyl RNA ASO, INSR ex2 caused IR exon 11 skipping by targeting the same nucleotides as the ineffective INSR PNA ex2 (18mer). Further it was also clear that 2′-O-methyl RNA ASOs exert their effect on IR exon 11 splicing at significantly lower concentrations than PNA-peptide hybrid ASOs. These findings may suggest two (not mutually exclusive) possibilities for the PNA only ASOs:

1. Failure of localisation – either due to poor transfection efficiency or failure to bind to the target.
2. Localisation, but failure to inhibit splicing of IR exon 11 - The PNA only ASOs predicted to be effective on the basis of identical targeting to effective 2′-O-methyl RNA ASOs, bind adequately to the require region but do not cause IR exon 11 skipping. This may be the result of failure to alter secondary structure or failure to block interaction with spliceosome/splicing factors.

4.1.3.2 PNA-peptide hybrid ASOs cause exon skipping in the context of the IR exon 11.

PNA-peptide hybrid ASOs designed to enhance exon inclusion by recruiting enhancing splicing factors paradoxically caused exon skipping in the context of the IR exon 11. The results of these experiments were unexpected and provide

caveats for the use of this technology but also insights relating to the splicing of the IR exon 11. Several explanations for this effect may be considered:

1. Steric interference - PNA-peptide hybrid ASOs sterically interfere with the interaction between the spliceosome and IR exon 11 to cause skipping.
2. Secondary structure alteration - PNA-peptide hybrid ASOs cause alteration in secondary structure which are detrimental to exon 11 splicing.
3. Inappropriate recruitment - The RS domain tail may incorrectly recruit or mislocalise splicing factors for the splicing reaction.
4. Sequestration - PNA-peptide hybrid ASOs do not localise to IR exon 11 but instead sequester enhancing splicing factors away from the IR exon 11, resulting in reduced exon 11 inclusion. Against this, RS domain tail transfected in isolation do not alter IR exon 11 splicing (Chapter 3.2).

4.1.3.3 PNA-peptide hybrid ASOs associated with 3′ tails cause higher degrees of IR exon 11 skipping than ASOs associated with 5′ tails.

This study noted that PNA-peptide hybrid ASOs with 3′ tails appeared to cause more exon skipping than equivalent ASOs with 5′ tails. This trend was observed for PNA-peptide hybrid ASOs targeting both halves of IR exon 11 though a statistical difference was only noted between INSR PNA ex1 3′ tail and INSR PNA ex1 5′ tail (unpaired t test $p=0.0142$).

4.1.4 2′-O-methyl ASO targeting multiple ESS in APOB exon 26.

APOB exon 26 is a long constitutive exon which is spliced despite multiple stronger pseudo splice sites orientated correctly within it (3′ splice site → 5′ splice site). It is unclear why the constitutive splice sites are chosen over certain high scoring pseudo sites. Presumed enhancer and silencer elements in introns and exons are thought to exist which strengthen/weaken given splice sites. Previous work has identified multiple exonic splicing silencers in APOB exon 26 in vitro splicing assays. How these sites interacted and what effects they had in vivo was unclear.

This study demonstrated that multiple cis-acting elements in a constitutive exon acted in a co-operative fashion to modulate and block aberrant splicing. Blocking multiple selected splice sites had an additive effect in causing either exon 26 skipping or alternative splicing with an alternative downstream 3′ splice site. The effects seen were minor however. This may be ascribed to poor transfection efficiency of ASOs to ESS sites in APOB exon 26. Sub-optimal targeting multiply redundant ESS, both in terms of number and combinations may also have played a role in the limited effect seen.

While the splice site strength of this cryptically activated 3′ splice site is higher than the native 3′ splice site it is unclear why a group of ESS approximately 7,000 nucleotides upstream might be involved in repressing a pseudo splice site that does not have the highest position score. We may speculate that disruption of exonic silencers allows more efficient aggregation of splicing factors around cryptic splice sites further downstream, thereby increasing the kinetics of spliceosome assembly and allowing minor isoforms to become apparent. The other possibility is that interference with ESSs allows silencing protein activity to become more influential at the constitutive 3′ splice site. Other methodological explanations include the possibility that pseudo splice sites nearer the ESS sites are selected

which would give significantly larger products which were not amplified efficiently by the PCR in comparison to the shorter transcripts, or that could not be distinguished from constitutively spliced transcripts by gel electrophoresis.

4.2 Support of findings from published data.

4.2.1. Identifying sites important in IR exon 11 splicing using 2′-O-methyl RNA ASOs.

The results from the current study may be reviewed in the context of the findings of Sen et al. (Sen et al. 2009). The Sen model was derived from earlier IR minigene studies undertaken from the same laboratory. Mutations, deletions and insertions were introduced into a minigene construct containing the IR exon 11 and adjacent exons and truncated introns. The effects on splicing were assessed in an in vivo cell culture system. These data are discussed in section 1.4.5. The data on regulatory sites within exon 11 formed the basis for work to elucidate the splicing factors important in the splicing of the IR exon 11. Using RNA affinity purification, SR proteins derived from Hela nuclear extract, bound to immobilised exon 11 RNA were analysed by gel electrophoresis. By identifying which bands were removed in the presence of mutated exon 11 sequence and considering their size, splicing factor identification was confirmed using splicing factors specific monoclonal antibodies. These studies identified SRp40, CUG-BP1 and SF2/ASF as the splicing factors that interact with the IR exon 11.

Figure 4.2.1.B shows the overlap between ASOs and splicing factor binding sites in the IR exon 11. The 2′-O-methyl RNA ASO INSR ex1 blocked the 5′ end of IR exon 11 (nucleotides 1-18) but appeared to have no effect on its splicing. Figure 4.2.1.A shows that INSR ex1 would completely block the putative binding site for the enhancer SRp20 (nucleotide 7-12 in IR exon11) and one nucleotide of the binding site of the inhibitory factor CUG-BP1 (nucleotide 18). Given this, it may be expected that INSR ex1 would cause exon skipping, which it did not. Only if blocking nucleotide 18 were crucial to the CUG-BP1 effect might it be argued that

simultaneous blocking of the enhancer and inhibitory effect of SRp20 and CUG-BP1 respectively, explains the neutral effect of INSR ex1. Against this is the finding that CUG-BP1 appears to bind strongly to the ESS in exon 11 even in the context of mutations to the ESS and did not elute from the exon template with 1 M salt (Sen et al. 2009) hinting that INSR ex1, binding to nucleotide 18, is unlikely to displace the binding of CUG-BP1.

INSR ex2 (binding nucleotide 19-36 of IR exon11) would block the binding of both the inhibitor, CUG-BP1 (nucleotides 18-26) and the enhancer, SF2/ASF (nucleotides 26-30). A neutral effect of INSR ex2 might be expected in view of this but INSR ex2 causes IR exon 11 skipping. This may suggest that SF2/ASF plays the dominant role over CUG-BP1 in this context. The findings of this study with respect to the effect of INSR ex2 are in keeping with the identification of a ESE at the 3′ end of IR exon 11 however (Kosaki et al. 1998; Sen et al. 2009). INSR ex3 (binding nucleotides 10-28) has a neutral effect on IR exon 11 splicing but would putatively interfere with all the splicing factors wholly (CUG-BP1) or in part (SRp20 and SF2/ASF). INSR ex1+2 would block the effects of all the splicing factors completely but in contrast to INSR ex3 causes significant exon 11 skipping. These two suppositions are in keeping with the findings of Sen et al. in that the difference in the effectiveness of INSR ex3 and INSR ex1+2 most likely relates to the fact that INSR ex1+2 completely blocks the enhancer splicing factors (SRp20 and SF2/ASF) while INSR ex3 only partly blocks their effects.

The ASO INSR 6/7 which targeted the 43 base nucleotide site identified by Kosaki et al. (Kosaki et al. 1998) as an intronic inhibitory region to exon 11 splicing did not alter splicing. This suggested either that the ASO bound to the target sequence but

not sufficiently to block its inhibitory effects or that the deletion in the minigenes inadvertently resulted in the creation of an inhibitory region with the apposed intronic regions and did not block a pre-existing inhibitory region.

Figure 4.2.1.A

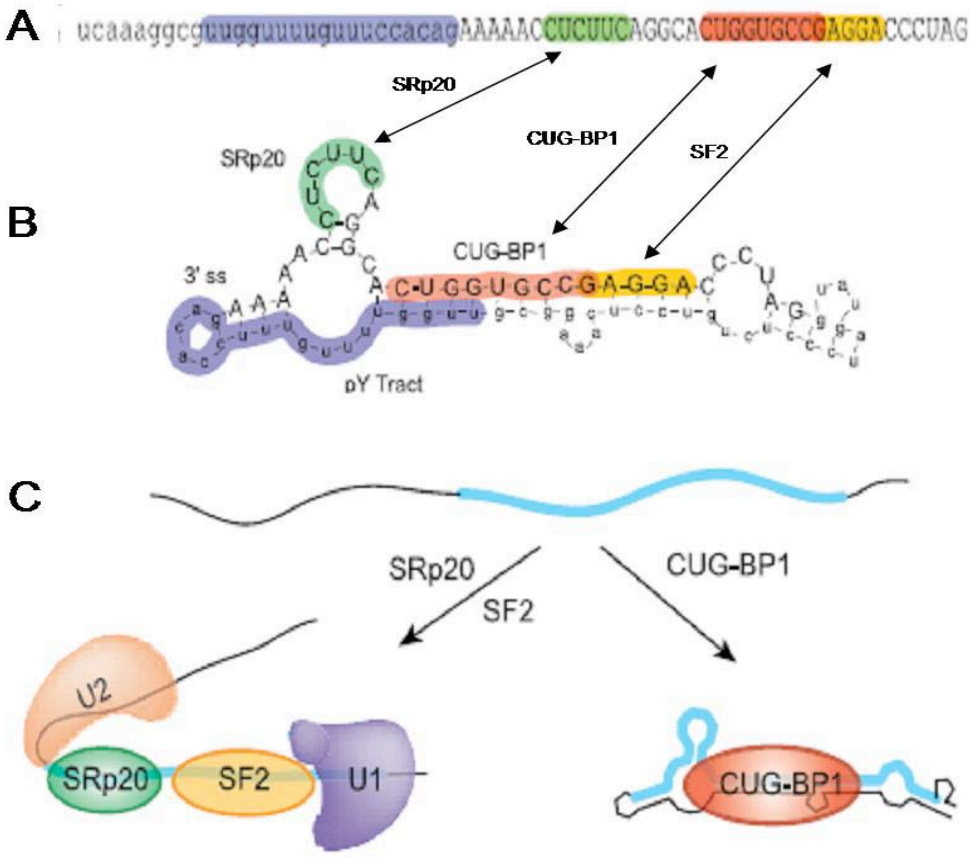


Figure 4.2.1.A Proposed model for interaction between IR exon 11 and splicing factors (modified from Sen et al. {Sen, 2009 #1070}). **A.** IR pre-mRNA showing exon 11 (capital letters) and putative binding sites for SRp20, CUG-BP1 and SF2/ASF. **B.** Secondary structure of IR pre-mRNA in the vicinity of exon 11 limits the binding of the spliceosome. C. Proposed model for splicing factor influence on IR exon 11 splicing. In the presence of SRp20 and SF2/ASF the conformation of pre-mRNA is favourable to spliceosome binding. In the presence of CUG-BP1 secondary structure conformation of the IR pre-mRNA is stabilised resulting in reduced spliceosome binding and IR exon 11 skipping.

Figure 4.2.1.B

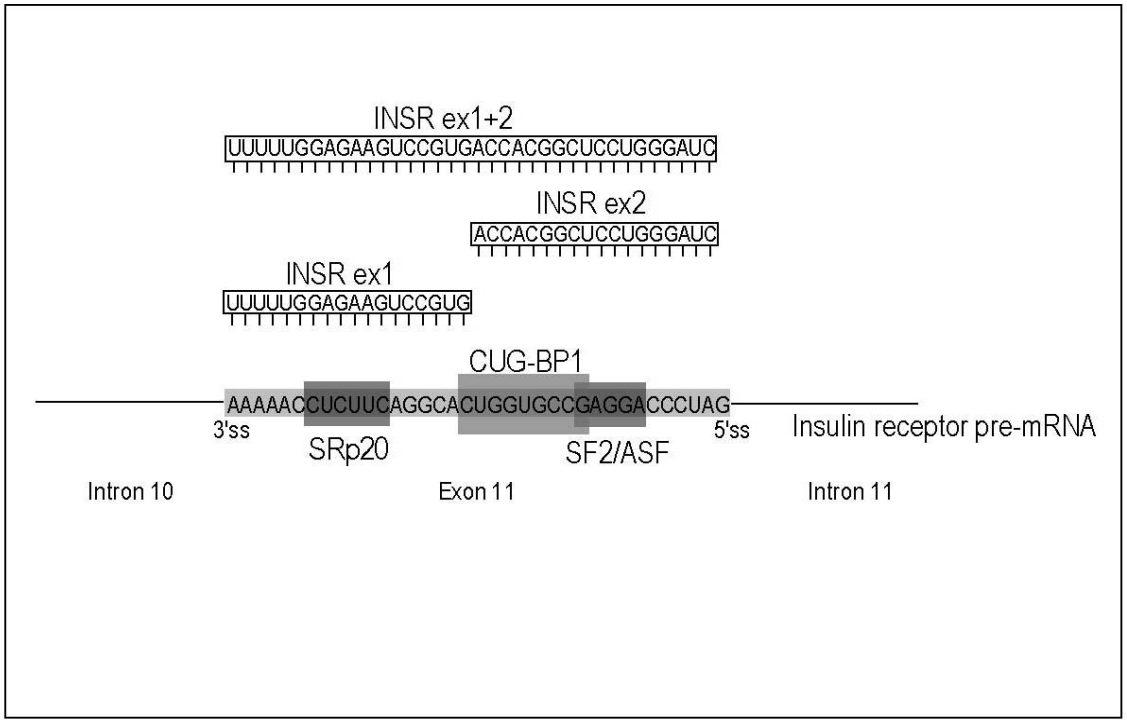


Figure 4.2.1.B Splicing factors and ASOs - interaction with the IR exon 11. Schematic diagram shows a portion of the IR pre-mRNA (not to scale). IR exon 11 and surrounding intronic sequences are shown. Exon 11 is given as the lettered sequence in the pre-mRNA while the introns are delineated by the grey lines. ASOs INSR ex1, INSR ex2 and INSR ex1+2 are illustrated above the IR pre-mRNA. Putative splicing factor interactions and their positions are given as grey boxes over exon 11. SRp20 binds at nucleotide position 7-12, CUG-BP1 binds at position 18-26 and SF2/ASF at nucleotide position 26-30. This figure demonstrates the overlap between ASO and splicing factor binding sites.

The minigene approach has the advantage over the ASO approach in studying splicing in that smaller and more specific changes can be assessed when investigating splicing choices. However the system is a poor physiological model. Over-expression of a minigene results in the alteration of the ratio of the concentration of splicing factors per transcript. Deletional studies may actually create artificial enhancer or silencer elements with the apposed ends of the transcript which are brought together instead of merely representing the effects of the deleted sequence. In the study by Kosaki et al. the ratio of the splice variants IR-B:IR-A was 60/40. This suggests that the splicing environment with the addition of the minigene differs from the native context. Our ASO strategy lacks some of the precision of the minigene approach but provides a more accurate physiological picture of splicing in vivo. Limits on how small the 2′-O-methyl RNA ASO can be made without compromising stability means that smaller regions of exon 11 could not be assessed. Regions of the exon may have silencer and enhancer properties which are not identified in this study because 18mer 2′-O-methyl RNA ASOs span both silencer and enhancer regions and nullify any observable effects. Hence the minigene approach provides a greater degree of precision in a less physiological model whilst ASOs provided less precision but in a more physiological model.

Work carried out by Khoo et al. (Khoo et al. 2007) transfecting 2′-O-methyl RNA ASOs targeting APOB exon 27, using a similar experimental set-up, may be viewed in comparison to work on the IR exon 11. In their experiments, intronic combination ASOs cause maximal skipping in the order ASO INSR 3B > ASO INSR 5B > ASO INSR 53. In terms of magnitude of effect, this is the opposite order to ASOs targeting IR exon 11 (INSR 53 > INSR 5B > INSR 3B). Though there were some differences between the studies it is of note that ASO effects were not consistent.

There do not appear to be any universal rules which can be applied to predicting the effects of ASOs on exon splicing.

4.2.2 TOES ASOs targeting the IR exon 11.

TOES ASOs did not alter IR exon 11 splicing. These compounds have been utilised effectively in other splicing systems using the same techniques but difference between the studies are noted and include (Skordis et al. 2003):

1. The use of 2′-O-methyl and phosphorothioate modifications for the ASO base as opposed to 2′-O-methyl RNA ASOs only. TOES ASOs with a 2′-O-methyl base were less effective at causing SMN2 exon 7 inclusion compared with those with 2′-O-methyl and phosphorothioate ASO base (Owen et al. 2011).
2. The transfection agent was jetPEI (Qbiogene, Nottingham, U.K.) as opposed to lipofectamine 2000 in this case.
3. The exon targeted for inclusion (SMN2 exon 7) was 54 nucleotides in length in comparison to the smaller 36 nucleotide exon 11 of the IR.
4. Both a 5′ and 3′ tail were used with the same ASO base in this study in comparison only a 5′ tail in the previous study.
5. TOES ASOs targeted a range of nucleotides in other studies on and around exon 7 of SMN2 whilst the first 18 nucleotides were targeted in this study including the nucleotide making the exon/intron border.

What had not been delineated clearly in the IR exon 11 model at the time when the TOES ASO design was undertaken was which splicing factors were important in exon 11 splicing and exactly where they bound. In the light of the work by Sen at al. (Sen et al. 2009) this now may be retrospectively reviewed. The ESE tail of the TOES compound has the repetitive sequence ‘GGA’ which binds the splicing factor SF2/ASF shown to bind IR exon 11. It does not carry motifs which might bind SRp20 or CUG-BP1. A recent study assessed different tail sequences in TOES ASOs. The tail used here with the ‘GGA’ motif tail conferred the highest efficacy in

the context of SMN2 exon 7. The data also suggest that matching tail sequences to splicing factors is not entirely clear given that alternative tails with sequences which should recruit SF2/ASF, did not (Owen et al. 2011).

The TOES ASO bound to nucleotides 4 - 18 of IR exon 11. The TOES ASO tested would have blocked the SRp20 binding site (nucleotides 7-12) (figure 4.2.1.B), allowed the majority of CUG-BP1 to bind (except over nucleotide 18) and presumably attract SF2/ASF appropriately but in the context of an unhindered SF2/ASF binding site (nucleotides 26-30). In addition, SF2/ASF is known to be expressed in abundance in HepG2 cells (Sen et al. 2009). The overall effect this compound may have added might therefore be negligible.

Other studies have demonstrated that 5' splice site utilisation may be altered by factors bound 18 to 25 nucleotides upstream of a 5' splice site (Chabot et al. 1987; Cunningham et al. 1991).

4.2.3 PNA and PNA-peptide hybrid ASOs targeting the IR exon 11 .

4.2.3.1 Equivalent PNA and 2′-O-methyl RNA ASOs do not have equivalent effect on splicing in the IR exon 11 splicing.

Comparison of the work using tailed ASOs revealed that the use of a 2′-O-methyl/phosphothoriate ASO base in a TOES ASO was associated with a higher efficacy in SMN2 exon 7 inclusion at lower concentrations than a PNA molecule in an ESSENCE compound (Cartegni et al. 2003; Skordis et al. 2003). However the value of this observation is limited given the differences between the ASOs; tail sequences; ASO binding sites and different experimental conditions. A more direct comparison, comes from a study comparing the effects of ASOs on the trans-activating response region (*TAR)* of the HIV genome which revealed a lower IC_{50}

for 2′-O-methyl RNA ASOs (20nM) compared with equivalent PNA ASOs (40nM) (Boulme et al. 1998). Failure to localise the PNA only ASOs may relate, in part, to their neutral backbone. While making these compounds extremely stable and highly avid for binding RNA it may make cell penetration more problematic. Since the PNA ASOs alone had no effect on splicing but the tailed PNA-peptide hybrid ASOs resulted in exon skipping, one may surmise that the effects of the charged tail (positively charged arginine due to the associated guanidinium group whilst serine is negatively charged due to the hydroxyl grouping) may facilitate transfection of these compounds in comparison to PNA ASOs alone.

4.2.3.2 PNA-peptide hybrid ASOs cause exon skipping in the context of the IR exon 11.

Previous uses of this technology have resulted in exon inclusion in the context of splicing models BRCA, SMN and BCL (Cartegni et al. 2003; Wilusz et al. 2005).

In the context of work by Sen et al. (Sen et al. 2009) it is clear the both PNA ASO targeting components of the PNA-peptide hybrid ASOs i.e. INSR PNA ex1 and INSR PNA ex2 targeted regions which bound both splicing enhancers (SRp20 and SF2/ASF) and inhibitors (CUG-BP1). The potential splicing enhancing properties of the PNA-peptide hybrid ASOs may have been negated by binding (and so blocking) natural splicing enhancer sites.

SR proteins such as SF2/ASF, which are normally thought of as enhancers of splicing, can act as splicing repressors when bound to intronic elements (Graveley 2000; Dauksaite et al. 2002). Exonic splicing enhancer sequences (ESE) can behave as splicing inhibitors when placed in an intron of a β-globin reporter system (Ibrahim el et al. 2005). It is not clear however, how SR proteins in these contexts cause splicing repression. These examples however involve SR proteins and ESE sequences in intronic regions and not RS domains as part of ASOs.

4.2.3.3 3′ tails associated with PNA-peptide hybrid ASOs caused higher degrees of IR exon 11 skipping than 5′ associated tails.

Previous studies use tails at both 3′ and 5′ ends of the ASO and all these compounds are able to cause exon inclusion however, none of these studies compare tails at opposing ends of ASOs. (Cartegni et al. 2003; Skordis et al. 2003; Wilusz et al. 2005). Two of these studies assess different tailed strategies (TOES and ESSENCE ASOs) and use tails at different ends of the molecules in an in vitro SMN2 assay system. Though complete comparison can not be made because of

different assay set-ups and ASO chemistries used, it is of note that the TOES ASOs with a 5′ tail, at approximately a tenth the concentration, result in double the effect on exon 7 inclusion in the SMN2 assay system in comparison to the ESSENCE ASO with a 3′ tail. Our data shows the opposite effect, but with respect to exon *exclusion* i.e. 3′ tails were twice as effective at causing exon exclusion compared to 5′ tails. This may therefore suggest that 5′ tail have some inherent advantage over 3′ tails in enhancing exon inclusion.

4.2.4 Constitutive APOB exon 26 can be skipped by inactivating multiple exonic splicing silencers within it.

In most experimentally characterised cases, ESS have been isolated as single elements or sometimes bipartite elements of variable length from 5 nucleotides to 119 nucleotides, usually in the context of alternative splicing (Caputi et al. 1994; Staffa et al. 1997). The current data, using 2′-O-methyl RNA ASOs to block these sites provides the first evidence that silencer elements appear to work in tandem to block pseudo splice sites in vivo in large constitutive exons such as APOB exon 26.

4.3 Experimental limitations.

The limitations of the current study relate to three principle questions.

1. **Confounding factors** - Did potential confounding factors account for some of the effects or lack of effects seen and attributed to ASOs?
2. **Specificity** - Were the effects of given ASOs specific to their intended targets?
3. **Quantification** - What was the magnitude of the effect and was it physiologically relevant?
4. **Relevance** - Were the ESSs targeted in APOB exon 26 appropriate to demonstrate a significant effect on splicing?

4.3.1. Confounding factors

It may be argued that the efficacy or lack of efficacy of the ASOs tested both in the context of splicing of the IR exon 11 and APOB exon 26 may have related to the effect of confounding factors. Variable degrees of transfection efficiency may account for differences in the apparent effects of the ASOs. Transfection efficiency was partially assessed in experiment section 3.1.4 but this was not quantified and all ASOs were not comprehensively assessed.

Cell viability is a consideration when assessing the effects of ASOs. It is possible that ASOs may affect cell viability differentially and this may have subtle knock-on effects on the isoform ratios seen. No obvious evidence of cell death as a result of transfecting different ASOs in different quantities was seen via an assessment of GAPDH and microscopy. However cell viability was not assessed in further detail in this study.

Secondary structure consideration may account either for the apparent lack of efficacy of some of the ASOs or the mechanism by which other ASOs exert their effects.

No TOES ASO or PNA-peptide hybrid ASO caused IR exon 11 inclusion as they were designed to do. A positive control included in these experiments was INSR ex2 to demonstrate transfection resulting in IR exon 11 skipping was still occurring. However the experiment lacked a TOES ASO or PNA-peptide hybrid ASO positive control. Demonstrating that a TOES ASO or PNA-peptide hybrid ASO could cause exon inclusion, in a different splicing model e.g. SMN2, using the current experimental set-up is an important further positive control. In its absence it could be argued that there may be a methodological impediment to causing exon inclusion rather that failure of the TOES ASO or PNA-peptide hybrid ASO.

4.3.2. Specificity

Control experiments were carried out to confirm the specificity of ASOs to their targets. These included transfection of scramble ASOs in the sample cell culture model to compare the effects on splicing (section 3.1.2). These control experiments did not confirm that ASOs only alter splicing in the target gene and not in other alternatively spliced genes. If the ASOs are specific, there should be no effect on the splicing of other alternatively spliced genes. These experiments have not done in relation to the INSR ASOs.

4.3.3. Quantification

Semi-quantitative PCR was used to quantify proportions of the IR isoforms and APOB exon 26 splicing products in this study. Semi-quantitative PCR was used throughout for consistency, because it had been used in comparable studies in the literature beforehand and it provided enough resolution for the purposes of this study. Real-time PCR would provide more precise quantitative data.

The effects of ASOs were expressed in terms of levels of RNA expression in this study. One way to clarify whether the ASOs caused a change in splicing which was biologically relevant would be confirm these effects at the protein level. Given that the effects on splicing were relatively minor in some cases e.g. effects of ASOs on APOB exon 26, it would be unlikely that these changes would be discriminated at the protein level.

4.3.4. Relevance

The aim of experiments in chapter 3.3 was to target predicted ESS sequences in APOB exon 26 separately and in tandem to see whether they de-repressed any alternative splice sites within exon 26. Multiple ESS where predicted along the length of APOB exon 26 however it was not clear how many sites interacted in tandem, the location of these sites and what was the target pseudo splice site might be. The apparent limited affect of the blocking multiple ESSs in APOB exon 26 may hide the potential importance of the mechanism if non-significant ESS sites were blocked.

4.4 Final conclusions.

The IR exon 11 represents a short, alternatively spliced exon which can easily be skipped using ASOs but in which exon inclusion can not be enhanced using the tested ASO technologies. APOB exon 26, in comparison, is a large, constitutively spliced exon, in which splicing can be slightly modified by targeting multiple sites with ASOs. Though the principles governing the splicing process and the use of antisense technology are consistent, study of these contrasting models emphasises that there is no universal strategy to alter splicing in these transcripts. We find that previously published data using 2′-O-methyl ASOs, TOES ASOs and PNA-peptide hybrid ASOs are not applicable to the endogenous IR exon 11. The actions of these trans-acting factors are not predictable from current knowledge. Rather, the effective use of this technology in each system requires an appreciation of the transcript itself, the unique composition of the splicing environment around it and the interaction between the two.

The use of ASOs is increasingly making the transition from laboratory to bedside. Drugs based on this technology have been commercially available for some time now with many others in development. ASOs provide the attractive proposition of providing designer drugs with a high specificity for their targets however there still remain factors limiting their efficacy. As investigative tools, though precise and for all their advantages, ASOs are unlikely to be able to spatially resolve short, closely apposed modifier sites whilst maintaining specificity to the target as illustrated in the case of the enhancer and silence regions identified in IR exon 11 by other methodologies. There remains a constant trade off with ASO targeting between precision and specificity.

The effects demonstrated in the study are relatively limited, however, as refinements occur, principally in ASO chemistry and drug delivery but also our

ability to appreciate and model the local splicing environment, these compounds will become even more powerful in the study and therapeutic manipulation of target transcripts and diseases as will our ability to deploy them.

4.5 Future work.

4.5.1. Confounding factors

Certain ASOs may have been ineffective as a result of poor transfection efficiency or poor binding to the target pre-mRNA. Transfection efficiency could be assessed more formally by cell counting using fluorescein-tagged ASOs and confocal microscopy (Chapter 3.1.4) or via fluorescence activated cell sorting (FACS).

If transfection efficiency is not found to be a significant issue then variability in binding to the pre-mRNA could be assessed by utilising a gel-shift binding assay. A pre-mRNA transcript including the IR exon 11 and radiolabelled ASOs would be subject to an in vitro splicing reaction. Free and bound ASO would be resolved using a polyacrylamide gel, excised and quantified. Binding affinity of the ASO would be proportional to the radioactivity count of the bound fraction. Whilst this method may provide some information regarding relative binding under assay conditions, it is unlikely to provide realistic insights into the binding of the ASO in the in vivo nuclear environement.

The lack of effect of some ASOs and the mechanism by which others exert their effect may be due to secondary structure considerations. It has been suggested that secondary structure variation around exon 11 could affect the binding of ASOs, splicing factors and spliceosome components to alter exon 11 splicing. Plasmids are available which contain either a wild type IR exon 11 pre-mRNA transcript or ones which contain various mutant sequences. These mutated sequences have been designed to be resistant to secondary structure formation whilst retaining the ability to bind ASOs. Co-transfecting these plasmids with ASOs in HepG2 cell

culture may provide evidence as to whether in the absence of secondary structure considerations ASOs become more or less effective.

To assess whether ASO effects result from inhibiting the binding of key trans-acting factors a co-immunoprecipitation experiment could be undertaken. ASOs including the PNA-peptide hybrid ASOs would be added to in vitro splicing reactions using radiolabelled IR pre-mRNA transcripts. Sepharose beads bound to antibodies against splicing factors known to interact with the IR (SRp20, SF2/ASF and CUG-BP1) or against U1 snRNP would be added to the splicing reaction followed by phenol extraction and ethanol precipitation. The amount of transcript bound would be estimated from the amount of radioactivity detected. If the given ASO inhibited the binding of a splicing factor or a U1 SnRNP then less radioactive pre-mRNA transcript would be detected.

The lack of efficacy of the TOES ASOs and PNA-peptide hybrids may be related to methodological errors which have not been identified. To confirm the experimental set-up is correct a further positive control would be crucial. TOES ASOs and PNA-peptide hybrids have been used to enhance exon 7 inclusion in SMN2 and experiment should been run concurrently with the previous experiments to demonstrate that exon inclusion is possible in the experimental set-up used.

4.5.2. Specificity

To further confirm the specificity of the ASOs targeting IR exon 11 other alternatively spliced genes e.g. TSC2 exon 25, Caspase 9 exon 3-6, could be examined from Hep G2 cells transfected with these ASOs to see whether they cause any off-target isoform changes. For greater coverage an exon-specific

microarray looking for alternative splicing events induced by ASOs e.g. Affymetrix Human Exon 1.0ST assay could be used on RNA harvested from transfected cells

4.5.3. Quantification

Real-time PCR will provide a more precise quantatative analysis of the effectiveness of a given ASO.

Western blotting could be employed to confirm that the levels of IR-A and IR-B protein are equivalent to the changes demonstrated with RNA. Differences in protein concentration are unlikely to be demonstrated in the context of APOB exon 26 however given the minor effect on splicing demonstrated by blocking ESS sites tested in exon 26.

4.5.4. Relevance

The results from Chapter 3.3 demonstrate that targeting multiple ESS in APOB exon 26 causes aberrant splicing of a constitutive exon to a minor degree. The magnitude of this effect may be explained by various factors including transfection efficiency and sub-optimal targeting of ESS sites. As an extension to this work it would be of value to assess the effects of targeting further combinations of sites both in close proximity and at a distance to one another to confirm whether splicing may be altered to a greater degree than the initial experiments suggest.

Were the ESSs targeted in APOB exon 26 appropriate to reveal the real silenced pseudoexons? An alternative strategy may be to attempt to switch splicing to favour expression of the 3 pseudoexons within APOB exon 26, calculated to have stronger combined splice sites scores (Shapiro & Senapathy) than the constitutive exon 26 splice sites. Confirming pseudoexon selection by targeting multiple neighbouring ESS sites around these combined splice sites would provide persuasive evidence that the mechanism is physiologically relevant.

Chapter 5

References

5.1 References.

Abelson, J. (2008). "Is the spliceosome a ribonucleoprotein enzyme?" Nat Struct Mol Biol 15(12): 1235-7.

Accili, D., J. Drago, et al. (1996). "Early neonatal death in mice homozygous for a null allele of the insulin receptor gene." Nat Genet 12(1): 106-9.

Akdim, F., E. S. Stroes, et al. (2007). "Antisense apolipoprotein B therapy: where do we stand?" Curr Opin Lipidol 18(4): 397-400.

Akdim, F., M. E. Visser, et al. (2010). "Effect of mipomersen, an apolipoprotein B synthesis inhibitor, on low-density lipoprotein cholesterol in patients with familial hypercholesterolemia." Am J Cardiol 105(10): 1413-9.

Akker, S. A., P. J. Smith, et al. (2001). "Nuclear post-transcriptional control of gene expression." J Mol Endocrinol 27(2): 123-31.

Alcazar-Roman, A. R., E. J. Tran, et al. (2006). "Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export." Nat Cell Biol 8(7): 711-6.

Allmang, C., E. Petfalski, et al. (1999). "The yeast exosome and human PM-Scl are related complexes of 3' --> 5' exonucleases." Genes Dev 13(16): 2148-58.

Altmann, K. H., D. Fabbro, et al. (1996). "Second-generation antisense oligonucleotides: structure-activity relationships and the design of improved signal-transduction inhibitors." Biochem Soc Trans 24(3): 630-7.

Amantana, A. and P. L. Iversen (2005). "Pharmacokinetics and biodistribution of phosphorodiamidate morpholino antisense oligomers." Curr Opin Pharmacol 5(5): 550-5.

Anderson, C. M., R. R. Henry, et al. (1993). "Relative expression of insulin receptor isoforms does not differ in lean, obese, and noninsulin-dependent diabetes mellitus subjects." J Clin Endocrinol Metab 76(5): 1380-2.

Ando, S., N. J. Sarlis, et al. (2001). "Aberrant alternative splicing of thyroid hormone receptor in a TSH-secreting pituitary tumor is a mechanism for hormone resistance." Mol Endocrinol 15(9): 1529-38.

Arenas, J. E. and J. N. Abelson (1997). "Prp43: An RNA helicase-like factor involved in spliceosome disassembly." Proc Natl Acad Sci U S A 94(22): 11798-802.

Ars, E., E. Serra, et al. (2000). "Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1." Hum Mol Genet 9(2): 237-47.

Ast, G. (2004). "How did alternative splicing evolve?" Nat Rev Genet 5(10): 773-82.

Auboeuf, D., A. Honig, et al. (2002). "Coordinate regulation of transcription and splicing by steroid receptor coregulators." Science 298(5592): 416-9.

Bai, Y., D. Lee, et al. (1999). "Control of 3' splice site choice in vivo by ASF/SF2 and hnRNP A1." Nucleic Acids Res 27(4): 1126-34.

Baitei, E. Y., M. Zou, et al. (2009). "Aberrant BRAF splicing as an alternative mechanism for oncogenic B-Raf activation in thyroid carcinoma." J Pathol 217(5): 707-15.

Baughan, T. D., A. Dickson, et al. (2009). "Delivery of bifunctional RNAs that target an intronic repressor and increase SMN levels in an animal model of spinal muscular atrophy." Hum Mol Genet 18(9): 1600-11.

Bauren, G. and L. Wieslander (1994). "Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription." Cell 76(1): 183-92.

Bejerano, G., M. Pheasant, et al. (2004). "Ultraconserved elements in the human genome." Science 304(5675): 1321-5.

Belfiore, A., G. Pandini, et al. (1999). "Insulin/IGF-I hybrid receptors play a major role in IGF-I signaling in thyroid cancer." Biochimie 81(4): 403-7.

Bellare, P., E. C. Small, et al. (2008). "A role for ubiquitin in the spliceosome assembly pathway." Nat Struct Mol Biol 15(5): 444-51.

Beltran, M., I. Puig, et al. (2008). "A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition." Genes Dev 22(6): 756-69.

Benecke, H., J. S. Flier, et al. (1992). "Alternatively spliced variants of the insulin receptor protein. Expression in normal and diabetic human tissues." J Clin Invest 89(6): 2066-70.

Benyoucef, S., K. H. Surinya, et al. (2007). "Characterization of insulin/IGF hybrid receptors: contributions of the insulin receptor L2 and Fn1 domains and the alternatively spliced exon 11 sequence to ligand binding and receptor activation." Biochem J 403(3): 603-13.

Berget, S. M. (1995). "Exon recognition in vertebrate splicing." J Biol Chem 270(6): 2411-4.

Berglund, J. A., K. Chua, et al. (1997). "The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC." Cell 89(5): 781-7.

Bessonov, S., M. Anokhina, et al. (2008). "Isolation of an active step I spliceosome and composition of its RNP core." Nature 452(7189): 846-50.

Black, D. L. (1991). "Does steric interference between splice sites block the splicing of a short c-src neuron-specific exon in non-neuronal cells?" Genes Dev 5(3): 389-402.

Blanchette, M. and B. Chabot (1999). "Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization." EMBO J 18(7): 1939-52.

Blanchette, M., R. E. Green, et al. (2005). "Global analysis of positive and negative pre-mRNA splicing regulators in Drosophila." Genes Dev 19(11): 1306-14.

Blencowe, B. J. (2000). "Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases." Trends Biochem Sci 25(3): 106-10.

Bonnal, S., C. Martinez, et al. (2008). "RBM5/Luca-15/H37 regulates Fas alternative splice site pairing after exon definition." Mol Cell 32(1): 81-95.

Bono, F., J. Ebert, et al. (2006). "The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA." Cell 126(4): 713-25.

Boucher, L., C. A. Ouzounis, et al. (2001). "A genome-wide survey of RS domain proteins." RNA 7(12): 1693-701.

Boulme, F., F. Freund, et al. (1998). "Modified (PNA, 2'-O-methyl and phosphoramidate) anti-TAR antisense oligonucleotides as strong and specific inhibitors of in vitro HIV-1 reverse transcription." Nucleic Acids Res 26(23): 5492-500.

Boutz, P. L., P. Stoilov, et al. (2007). "A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons." Genes Dev 21(13): 1636-52.

Bracco, L. and J. Kearsey (2003). "The relevance of alternative RNA splicing to pharmacogenomics." Trends Biotechnol 21(8): 346-53.

Brinster, R. L., J. M. Allen, et al. (1988). "Introns increase transcriptional efficiency in transgenic mice." Proc Natl Acad Sci U S A 85(3): 836-40.

Buettner, C., R. Patel, et al. (2005). "Severe impairment in liver insulin signaling fails to alter hepatic insulin action in conscious mice." J Clin Invest 115(5): 1306-13.

Buratti, E., A. F. Muro, et al. (2004). "RNA folding affects the recruitment of SR proteins by mouse and human polypurinic enhancer elements in the fibronectin EDA exon." Mol Cell Biol 24(3): 1387-400.

Burge, C. and S. Karlin (1997). "Prediction of complete gene structures in human genomic DNA." J Mol Biol 268(1): 78-94.

Burgess, S. M. and C. Guthrie (1993). "A mechanism to enhance mRNA splicing fidelity: the RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates." Cell 73(7): 1377-91.

Butler, M. G. (2011). "Prader-Willi Syndrome: Obesity due to Genomic Imprinting." Curr Genomics 12(3): 204-15.

Caceres, J. F., G. R. Screaton, et al. (1998). "A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm." Genes Dev 12(1): 55-66.

Caputi, M., G. Casari, et al. (1994). "A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon." Nucleic Acids Res 22(6): 1018-22.

Carmody, S. R. and S. R. Wente (2009). "mRNA nuclear export at a glance." J Cell Sci 122(Pt 12): 1933-7.

Cartegni, L., S. L. Chew, et al. (2002). "Listening to silence and understanding nonsense: exonic mutations that affect splicing." Nat Rev Genet 3(4): 285-98.

Cartegni, L. and A. R. Krainer (2003). "Correction of disease-associated exon skipping by synthetic exon-specific activators." Nat Struct Biol 10(2): 120-5.

Castillo-Davis, C. I., S. L. Mekhedov, et al. (2002). "Selection for short introns in highly expressed genes." Nat Genet 31(4): 415-8.

Chabot, B. and J. A. Steitz (1987). "Multiple interactions between the splicing substrate and small nuclear ribonucleoproteins in spliceosomes." Mol Cell Biol 7(1): 281-93.

Chan, L. (1992). "Apolipoprotein B, the major protein component of triglyceride-rich and low density lipoproteins." J Biol Chem 267(36): 25621-4.

Chan, S. P., D. I. Kao, et al. (2003). "The Prp19p-associated complex in spliceosome activation." Science 302(5643): 279-82.

Chang, Y. F., J. S. Imam, et al. (2007). "The nonsense-mediated decay RNA surveillance pathway." Annu Rev Biochem 76: 51-74.

Charlet, B. N., R. S. Savkur, et al. (2002). "Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing." Mol Cell 10(1): 45-53.

Chastain, M. and I. Tinoco, Jr. (1991). "Structural elements in RNA." Prog Nucleic Acid Res Mol Biol 41: 131-77.

Cheah, M. T., A. Wachter, et al. (2007). "Control of alternative RNA splicing and gene expression by eukaryotic riboswitches." Nature 447(7143): 497-500.

Chen, C. D., R. Kobayashi, et al. (1999). "Binding of hnRNP H to an exonic splicing silencer is involved in the regulation of alternative splicing of the rat beta-tropomyosin gene." Genes Dev 13(5): 593-606.

Chen, C. Y. and A. B. Shyu (2003). "Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway." Mol Cell Biol 23(14): 4805-13.

Chen, I. T. and L. A. Chasin (1994). "Large exon size does not limit splicing in vivo." Mol Cell Biol 14(3): 2140-6.

Chen, M. and J. L. Manley (2009). "Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches." Nat Rev Mol Cell Biol 10(11): 741-54.

Cheng, H., K. Dufu, et al. (2006). "Human mRNA export machinery recruited to the 5' end of mRNA." Cell 127(7): 1389-400.

Chester, A., J. Scott, et al. (2000). "RNA editing: cytidine to uridine conversion in apolipoprotein B mRNA." Biochim Biophys Acta 1494(1-2): 1-13.

Chew, S. L., P. Lavender, et al. (1995). "An alternatively spliced human insulin-like growth factor-I transcript with hepatic tissue expression that diverts away from the mitogenic IBE1 peptide." Endocrinology 136(5): 1939-44.

Chia, S., S. Dent, et al. (2009). "Phase II trial of OGX-011 in combination with docetaxel in metastatic breast cancer." Clin Cancer Res 15(2): 708-13.

Chou, M. Y., N. Rooke, et al. (1999). "hnRNP H is a component of a splicing enhancer complex that activates a c-src alternative exon in neuronal cells." Mol Cell Biol 19(1): 69-77.

Company, M., J. Arenas, et al. (1991). "Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes." Nature 349(6309): 487-93.

Corradini, R., S. Sforza, et al. (2011). "Peptide nucleic acids with a structurally biased backbone. Updated review and emerging challenges." Curr Top Med Chem 11(12): 1535-54.

Crooke, R. M., M. J. Graham, et al. (2005). "An apolipoprotein B antisense oligonucleotide lowers LDL cholesterol in hyperlipidemic mice without causing hepatic steatosis." J Lipid Res 46(5): 872-84.

Cunningham, S. A., A. J. Else, et al. (1991). "Influences of separation and adjacent sequences on the use of alternative 5' splice sites." J Mol Biol 217(2): 265-81.

Das, R., Z. Zhou, et al. (2000). "Functional association of U2 snRNP with the ATP-independent spliceosomal complex E." Mol Cell 5(5): 779-87.

Das, R., J. Yu, et al. (2007). "SR proteins function in coupling RNAP II transcription to pre-mRNA splicing." Mol Cell 26(6): 867-81.

Dauksaite, V. and G. Akusjarvi (2002). "Human splicing factor ASF/SF2 encodes for a repressor domain required for its inhibitory activity on pre-mRNA splicing." J Biol Chem 277(15): 12579-86.

de la Mata, M., C. R. Alonso, et al. (2003). "A slow RNA polymerase II affects alternative splicing in vivo." Mol Cell 12(2): 525-32.

de la Mata, M. and A. R. Kornblihtt (2006). "RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20." Nat Struct Mol Biol 13(11): 973-80.

Deery, E. C., E. N. Vithana, et al. (2002). "Disease mechanism for retinitis pigmentosa (RP11) caused by mutations in the splicing factor gene PRPF31." Hum Mol Genet 11(25): 3209-19.

Del Gatto-Konczak, F., M. Olive, et al. (1999). "hnRNP A1 recruited to an exon in vivo can function as an exon splicing silencer." Mol Cell Biol 19(1): 251-60.

Denis, M. M., N. D. Tolley, et al. (2005). "Escaping the nuclear confines: signal-dependent pre-mRNA splicing in anucleate platelets." Cell 122(3): 379-91.

Denley, A., E. R. Bonython, et al. (2004). "Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR." Mol Endocrinol 18(10): 2502-12.

Dietz, H. C., D. Valle, et al. (1993). "The skipping of constitutive exons in vivo induced by nonsense mutations." Science 259(5095): 680-3.

Doma, M. K. and R. Parker (2006). "Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation." Nature 440(7083): 561-4.

Dominski, Z. and R. Kole (1991). "Selection of splice sites in pre-mRNAs with short internal exons." Mol Cell Biol 11(12): 6075-83.

Dominski, Z. and R. Kole (1992). "Cooperation of pre-mRNA sequence elements in splice site selection." Mol Cell Biol 12(5): 2108-14.

Dredge, B. K., G. Stefani, et al. (2005). "Nova autoregulation reveals dual functions in neuronal splicing." EMBO J 24(8): 1608-20.

Du, H. and M. Rosbash (2002). "The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing." Nature 419(6902): 86-90.

Dye, M. J., N. Gromak, et al. (2006). "Exon tethering in transcription by RNA polymerase II." Mol Cell 21(6): 849-59.

Dziembowski, A., A. P. Ventura, et al. (2004). "Proteomic analysis identifies a new complex required for nuclear pre-mRNA retention and splicing." EMBO J 23(24): 4847-56.

Dziembowski, A., E. Lorentzen, et al. (2007). "A single subunit, Dis3, is essentially responsible for yeast exosome core activity." Nat Struct Mol Biol 14(1): 15-22.

Ebina, Y., M. Edery, et al. (1985). "Expression of a functional human insulin receptor from a cloned cDNA in Chinese hamster ovary cells." Proc Natl Acad Sci U S A 82(23): 8014-8.

Ebina, Y., L. Ellis, et al. (1985). "The human insulin receptor cDNA: the structural basis for hormone-activated transmembrane signalling." Cell 40(4): 747-58.

Eckstein, F. (2000). "Phosphorothioate oligodeoxynucleotides: what is their origin and what is unique about them?" Antisense Nucleic Acid Drug Dev 10(2): 117-21.

Egholm, M., O. Buchardt, et al. (1993). "PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules." Nature 365(6446): 566-8.

Entingh, A. J., C. M. Taniguchi, et al. (2003). "Bi-directional regulation of brown fat adipogenesis by the insulin receptor." J Biol Chem 278(35): 33377-83.

Eperon, L. P., I. R. Graham, et al. (1988). "Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase?" Cell 54(3): 393-401.

Fairbrother, W. G. and L. A. Chasin (2000). "Human genomic sequences that inhibit splicing." Mol Cell Biol 20(18): 6816-25.

Fairbrother, W. G., R. F. Yeh, et al. (2002). "Predictive identification of exonic splicing enhancers in human genes." Science 297(5583): 1007-13.

Feng, Y., M. T. Valley, et al. (2009). "SRp38 regulates alternative splicing and is required for Ca(2+) handling in the embryonic heart." Dev Cell 16(4): 528-38.

Fenger-Gron, M., C. Fillman, et al. (2005). "Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping." Mol Cell 20(6): 905-15.

Fisher, S. J. and C. R. Kahn (2003). "Insulin signaling is required for insulin's direct and indirect action on hepatic glucose production." J Clin Invest 111(4): 463-8.

Fong, Y. W. and Q. Zhou (2001). "Stimulatory effect of splicing factors on transcriptional elongation." Nature 414(6866): 929-33.

Fox-Walsh, K. L., Y. Dou, et al. (2005). "The architecture of pre-mRNAs affects mechanisms of splice-site pairing." Proc Natl Acad Sci U S A 102(45): 16176-81.

Frasca, F., G. Pandini, et al. (1999). "Insulin receptor isoform A, a newly recognized, high-affinity insulin-like growth factor II receptor in fetal and cancer cells." Mol Cell Biol 19(5): 3278-88.

Frattali, A. L., J. L. Treadway, et al. (1992). "Transmembrane signaling by the human insulin receptor kinase. Relationship between intramolecular beta subunit trans- and cis-autophosphorylation and substrate kinase activation." J Biol Chem 267(27): 19521-8.

Freund, M., C. Asang, et al. (2003). "A novel approach to describe a U1 snRNA binding site." Nucleic Acids Res 31(23): 6963-75.

Fruman, D. A., R. E. Meyers, et al. (1998). "Phosphoinositide kinases." Annu Rev Biochem 67: 481-507.

Fukumaki, Y., P. K. Ghosh, et al. (1982). "Abnormally spliced messenger RNA in erythroid cells from patients with beta+ thalassemia and monkey cells expressing a cloned beta+-thalassemic gene." Cell 28(3): 585-93.

Galy, V., O. Gadal, et al. (2004). "Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1." Cell 116(1): 63-73.

Ghigna, C., M. De Toledo, et al. (2010). "Pro-metastatic splicing of Ron proto-oncogene mRNA can be reversed: therapeutic potential of bifunctional oligonucleotides and indole derivatives." RNA Biol 7(4): 495-503.

Giovannone, B., M. L. Scaldaferri, et al. (2000). "Insulin receptor substrate (IRS) transduction system: distinct and overlapping signaling potential." Diabetes Metab Res Rev 16(6): 434-41.

Goemans, N. M., M. Tulinius, et al. (2011). "Systemic administration of PRO051 in Duchenne's muscular dystrophy." N Engl J Med 364(16): 1513-22.

Gonzalez-Porta, M., M. Calvo, et al. (2012). "Estimation of alternative splicing variability in human populations." Genome Res 22(3): 528-38.

Gornemann, J., K. M. Kotovic, et al. (2005). "Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex." Mol Cell 19(1): 53-63.

Gozani, O., J. Potashkin, et al. (1998). "A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site." Mol Cell Biol 18(8): 4752-60.

Graveley, B. R. (2000). "Sorting out the complexity of SR protein functions." Rna 6(9): 1197-211.

Graveley, B. R., K. J. Hertel, et al. (2001). "The role of U2AF35 and U2AF65 in enhancer-dependent splicing." RNA 7(6): 806-18.

Graveley, B. R. (2005). "Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures." Cell 123(1): 65-73.

Gruter, P., C. Tabernero, et al. (1998). "TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus." Mol Cell 1(5): 649-59.

Gudikote, J. P., J. S. Imam, et al. (2005). "RNA splicing promotes translation and RNA surveillance." Nat Struct Mol Biol 12(9): 801-9.

Guiraud-Dogan, C., A. Huguet, et al. (2007). "DM1 CTG expansions affect insulin receptor isoforms expression in various tissues of transgenic mice." Biochim Biophys Acta 1772(11-12): 1183-91.

Halligan, D. L. and P. D. Keightley (2006). "Ubiquitous selective constraints in the Drosophila genome revealed by a genome-wide interspecies comparison." Genome Res 16(7): 875-84.

Hansen, T., C. Bjorbaek, et al. (1993). "Expression of insulin receptor spliced variants and their functional correlates in muscle from patients with non-insulin-dependent diabetes mellitus." J Clin Endocrinol Metab 77(6): 1500-5.

Hastings, M. L., C. Milcarek, et al. (1997). "Expression of the thyroid hormone receptor gene, erbAalpha, in B lymphocytes: alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels." Nucleic Acids Res 25(21): 4296-300.

Hatton, A. R., V. Subramaniam, et al. (1998). "Generation of alternative Ultrabithorax isoforms and stepwise removal of a large intron by resplicing at exon-exon junctions." Mol Cell 2(6): 787-96.

Hawkins, J. D. (1988). "A survey on intron and exon lengths." Nucleic Acids Res 16(21): 9893-908.

Hentze, M. W. and A. E. Kulozik (1999). "A perfect message: RNA surveillance and nonsense-mediated decay." Cell 96(3): 307-10.

Hermann, T. and D. J. Patel (1999). "Stitching together RNA tertiary architectures." J Mol Biol 294(4): 829-49.

Hiller, M., Z. Zhang, et al. (2007). "Pre-mRNA secondary structures influence exon recognition." PLoS Genet 3(11): e204.

Ho, T. H., B. N. Charlet, et al. (2004). "Muscleblind proteins regulate alternative splicing." Embo J 23(15): 3103-12.

Hoskins, A. A., L. J. Friedman, et al. (2011). "Ordered and dynamic assembly of single spliceosomes." Science 331(6022): 1289-95.

House, A. E. and K. W. Lynch (2006). "An exonic splicing silencer represses spliceosome assembly after ATP-dependent exon recognition." Nat Struct Mol Biol 13(10): 937-44.

Hribal, M. L., L. Perego, et al. (2003). "Chronic hyperglycemia impairs insulin secretion by affecting insulin receptor expression, splicing, and signaling in RIN beta cell line and human islets of Langerhans." Faseb J 17(10): 1340-2.

Hua, Y., T. A. Vickers, et al. (2007). "Enhancement of SMN2 exon 7 inclusion by antisense oligonucleotides targeting the exon." PLoS Biol 5(4): e73.

Hua, Y., K. Sahashi, et al. (2011). "Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model." Nature 478(7367): 123-6.

Huang, L. S., S. C. Bock, et al. (1985). "Human apolipoprotein B cDNA clone isolation and demonstration that liver apolipoprotein B mRNA is 22 kilobases in length." Proc Natl Acad Sci U S A 82(20): 6825-9.

Huang, Y. and J. A. Steitz (2005). "SRprises along a messenger's journey." Mol Cell 17(5): 613-5.

Huang, Z., N. L. Bodkin, et al. (1994). "Hyperinsulinemia is associated with altered insulin receptor mRNA splicing in muscle of the spontaneously obese diabetic rhesus monkey." J Clin Invest 94(3): 1289-96.

Huang, Z., N. L. Bodkin, et al. (1996). "Altered insulin receptor messenger ribonucleic acid splicing in liver is associated with deterioration of glucose tolerance in the spontaneously obese and diabetic rhesus monkey: analysis of controversy between monkey and human studies." J Clin Endocrinol Metab 81(4): 1552-6.

Huntzinger, E., I. Kashima, et al. (2008). "SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan." RNA 14(12): 2609-17.

Ibrahim el, C., T. D. Schaal, et al. (2005). "Serine/arginine-rich protein-dependent suppression of exon skipping by exonic splicing enhancers." Proc Natl Acad Sci U S A 102(14): 5002-7.

International Human Genome Sequencing Consortium (2004). "Finishing the euchromatic sequence of the human genome." Nature 431(7011): 931-45.

Isken, O., Y. K. Kim, et al. (2008). "Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay." Cell 133(2): 314-27.

Jani, D., S. Lutz, et al. (2012). "Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export." Nucleic Acids Res.

Javitt, N. B. (1990). "Hep G2 cells as a resource for metabolic studies: lipoprotein, cholesterol, and bile acids." FASEB J 4(2): 161-8.

Jensen, K. K., H. Orum, et al. (1997). "Kinetics for hybridization of peptide nucleic acids (PNA) with DNA and RNA studied with the BIAcore technique." Biochemistry 36(16): 5072-7.

Jialal, I. (1998). "Evolving lipoprotein risk factors: lipoprotein(a) and oxidized low-density lipoprotein." Clin Chem 44(8 Pt 2): 1827-32.

Juge, F., C. Fernando, et al. (2010). "The SR protein B52/SRp55 is required for DNA topoisomerase I recruitment to chromatin, mRNA release and transcription shutdown." PLoS Genet 6(9).

Jurica, M. S. and M. J. Moore (2003). "Pre-mRNA splicing: awash in a sea of proteins." Mol Cell 12(1): 5-14.

Kadener, S., P. Cramer, et al. (2001). "Antagonistic effects of T-Ag and VP16 reveal a role for RNA pol II elongation on alternative splicing." EMBO J 20(20): 5759-68.

Kaida, D., M. G. Berg, et al. (2010). "U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation." Nature 468(7324): 664-8.

Kalli, K. R., O. I. Falowo, et al. (2002). "Functional insulin receptors on human epithelial ovarian carcinoma cells: implications for IGF-II mitogenic signaling." Endocrinology 143(9): 3259-67.

Kan, J. L. and M. R. Green (1999). "Pre-mRNA splicing of IgM exons M1 and M2 is directed by a juxtaposed splicing enhancer and inhibitor." Genes Dev 13(4): 462-71.

Kaneko, S., O. Rozenblatt-Rosen, et al. (2007). "The multifunctional protein p54nrb/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3' processing and transcription termination." Genes Dev 21(14): 1779-89.

Kashima, I., A. Yamashita, et al. (2006). "Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay." Genes Dev 20(3): 355-67.

Kellerer, M., R. Lammers, et al. (1992). "Distinct alpha-subunit structures of human insulin receptor A and B variants determine differences in tyrosine kinase activities." Biochemistry 31(19): 4588-96.

Kellerer, M., G. Sesti, et al. (1993). "Altered pattern of insulin receptor isotypes in skeletal muscle membranes of type 2 (non-insulin-dependent) diabetic subjects." Diabetologia 36(7): 628-32.

Khoo, B., X. Roca, et al. (2007). "Antisense oligonucleotide-induced alternative splicing of the APOB mRNA generates a novel isoform of APOB." BMC Mol Biol 8: 3.

Khoo, B. and A. R. Krainer (2009). "Splicing therapeutics in SMN2 and APOB." Curr Opin Mol Ther 11(2): 108-15.

Kielkopf, C. L., N. A. Rodionova, et al. (2001). "A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer." Cell 106(5): 595-605.

Kim, H., R. Klein, et al. (2004). "Estimating rates of alternative splicing in mammals and invertebrates." Nat Genet 36(9): 915-6; author reply 916-7.

Kim, M., N. J. Krogan, et al. (2004). "The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II." Nature 432(7016): 517-22.

Kishore, S. and S. Stamm (2006). "The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C." Science 311(5758): 230-2.

Kochetov, A. V. (2008). "Alternative translation start sites and hidden coding potential of eukaryotic mRNAs." Bioessays 30(7): 683-91.

Kohler, A. and E. Hurt (2007). "Exporting RNA from the nucleus to the cytoplasm." Nat Rev Mol Cell Biol 8(10): 761-73.

Kolasinska-Zwierz, P., T. Down, et al. (2009). "Differential chromatin marking of introns and expressed exons by H3K36me3." Nat Genet 41(3): 376-81.

Konarska, M. M., J. Vilardell, et al. (2006). "Repositioning of the reaction intermediate within the catalytic center of the spliceosome." Mol Cell 21(4): 543-53.

Kondo, T., D. Vicent, et al. (2003). "Knockout of insulin and IGF-1 receptors on vascular endothelial cells protects against retinal neovascularization." J Clin Invest 111(12): 1835-42.

Koodathingal, P., T. Novak, et al. (2010). "The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing." Mol Cell 39(3): 385-95.

Kosaki, A. and N. J. Webster (1993). "Effect of dexamethasone on the alternative splicing of the insulin receptor mRNA and insulin action in HepG2 hepatoma cells." J Biol Chem 268(29): 21990-6.

Kosaki, A., T. S. Pillay, et al. (1995). "The B isoform of the insulin receptor signals more efficiently than the A isoform in HepG2 cells." J Biol Chem 270(35): 20816-23.

Kosaki, A., J. Nelson, et al. (1998). "Identification of intron and exon sequences involved in alternative splicing of insulin receptor pre-mRNA." J Biol Chem 273(17): 10331-7.

Kotani, K., W. Ogawa, et al. (1998). "Requirement of atypical protein kinase clambda for insulin stimulation of glucose uptake but not for Akt activation in 3T3-L1 adipocytes." Mol Cell Biol 18(12): 6971-82.

Krecic, A. M. and M. S. Swanson (1999). "hnRNP complexes: composition, structure, and function." Curr Opin Cell Biol 11(3): 363-71.

Kuhn, A. N., M. A. van Santen, et al. (2009). "Stalling of spliceosome assembly at distinct stages by small-molecule inhibitors of protein acetylation and deacetylation." RNA 15(1): 153-75.

Kulkarni, R. N., J. C. Bruning, et al. (1999). "Tissue-specific knockout of the insulin receptor in pancreatic beta cells creates an insulin secretory defect similar to that in type 2 diabetes." Cell 96(3): 329-39.

Kurreck, J., E. Wyszko, et al. (2002). "Design of antisense oligonucleotides stabilized by locked nucleic acids." Nucleic Acids Res 30(9): 1911-8.

Kuyumcu-Martinez, N. M., G. S. Wang, et al. (2007). "Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-mediated hyperphosphorylation." Mol Cell 28(1): 68-78.

Kyburz, A., A. Friedlein, et al. (2006). "Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3' end processing and splicing." Mol Cell 23(2): 195-205.

Lai, M. C., R. I. Lin, et al. (2000). "A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins." J Biol Chem 275(11): 7950-7.

Lallena, M. J., K. J. Chalmers, et al. (2002). "Splicing regulation at the second catalytic step by Sex-lethal involves 3' splice site recognition by SPF45." Cell 109(3): 285-96.

Lareau, L. F., M. Inada, et al. (2007). "Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements." Nature 446(7138): 926-9.

Laskin, J. J., G. Nicholas, et al. "Phase I/II Trial of Custirsen (OGX-011), an Inhibitor of Clusterin, in Combination with a Gemcitabine and Platinum Regimen in Patients with Previously Untreated Advanced Non-small Cell Lung Cancer." J Thorac Oncol.

Lauro, D., Y. Kido, et al. (1998). "Impaired glucose tolerance in mice with a targeted impairment of insulin action in muscle and adipose tissue." Nat Genet 20(3): 294-8.

Le Hir, H., A. Nott, et al. (2003). "How introns influence and enhance eukaryotic gene expression." Trends Biochem Sci 28(4): 215-20.

Lejeune, F., X. Li, et al. (2003). "Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities." Mol Cell 12(3): 675-87.

Li, W. M., T. Barnes, et al. (2009). "Endoribonucleases--enzymes gaining spotlight in mRNA metabolism." FEBS J 277(3): 627-41.

Lin, S., G. Coutinho-Mansfield, et al. (2008). "The splicing factor SC35 has an active role in transcriptional elongation." Nat Struct Mol Biol 15(8): 819-26.

Listerman, I., A. S. Bledau, et al. (2007). "Extragenic accumulation of RNA polymerase II enhances transcription by RNA polymerase III." PLoS Genet 3(11): e212.

Liu, H. X., M. Zhang, et al. (1998). "Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins." Genes Dev 12(13): 1998-2012.

Liu, J., M. A. Carmell, et al. (2004). "Argonaute2 is the catalytic engine of mammalian RNAi." Science 305(5689): 1437-41.

Liu, X., Q. Jiang, et al. (2002). "Partial correction of endogenous DeltaF508 CFTR in human cystic fibrosis airway epithelia by spliceosome-mediated RNA trans-splicing." Nat Biotechnol 20(1): 47-52.

Luco, R. F., Q. Pan, et al. (2010). "Regulation of alternative splicing by histone modifications." Science 327(5968): 996-1000.

Luna, R., C. Gonzalez-Aguilera, et al. (2009). "Transcription at the proximity of the nuclear pore: a role for the THP1-SAC3-SUS1-CDC31 (THSC) complex." RNA Biol 6(2): 145-8.

Lund, M. K. and C. Guthrie (2005). "The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim." Mol Cell 20(4): 645-51.

Lynch, M. and J. S. Conery (2000). "The evolutionary fate and consequences of duplicate genes." Science 290(5494): 1151-5.

Mackereth, C. D., T. Madl, et al. (2011). "Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF." Nature 475(7356): 408-11.

Makarov, E. M., O. V. Makarova, et al. (2002). "Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome." Science 298(5601): 2205-8.

Makarova, O. V., E. M. Makarov, et al. (2004). "A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing." EMBO J 23(12): 2381-91.

Mandel, C. R., S. Kaneko, et al. (2006). "Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease." Nature 444(7121): 953-6.

Mansfield, S. G., H. Chao, et al. (2004). "RNA repair using spliceosome-mediated RNA trans-splicing." Trends Mol Med 10(6): 263-8.

Maquat, L. E. (2004). "Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics." Nat Rev Mol Cell Biol 5(2): 89-99.

Maroney, P. A., C. M. Romfo, et al. (2000). "Functional recognition of 5' splice site by U4/U6.U5 tri-snRNP defines a novel ATP-dependent step in early spliceosome assembly." Mol Cell 6(2): 317-28.

Massenet, S., L. Pellizzoni, et al. (2002). "The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway." Mol Cell Biol 22(18): 6533-41.

Mathew, R., K. Hartmuth, et al. (2008). "Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome." Nat Struct Mol Biol 15(5): 435-43.

Matter, N., P. Herrlich, et al. (2002). "Signal-dependent regulation of splicing via phosphorylation of Sam68." Nature 420(6916): 691-5.

Mayeda, A. and A. R. Krainer (1992). "Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2." Cell 68(2): 365-75.

Mayeda, A., D. M. Helfman, et al. (1993). "Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF." Mol Cell Biol 13(5): 2993-3001.

Mayes, A. E., L. Verdone, et al. (1999). "Characterization of Sm-like proteins in yeast and their association with U6 snRNA." EMBO J 18(15): 4321-31.

McClain, D. A. (1991). "Different ligand affinities of the two human insulin receptor splice variants are reflected in parallel changes in sensitivity for insulin action." Mol Endocrinol 5(5): 734-9.

McGlincy, N. J. and C. W. Smith (2008). "Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense?" Trends Biochem Sci 33(8): 385-93.

Mendell, J. T. and H. C. Dietz (2001). "When the message goes awry: disease-producing mutations that influence mRNA content and performance." Cell 107(4): 411-4.

Metherell, L. A., S. A. Akker, et al. (2001). "Pseudoexon activation as a novel mechanism for disease resulting in atypical growth-hormone insensitivity." Am J Hum Genet 69(3): 641-6.

Meyer, K., J. Marquis, et al. (2009). "Rescue of a severe mouse model for spinal muscular atrophy by U7 snRNA-mediated splicing modulation." Hum Mol Genet 18(3): 546-55.

Minn, A. H., M. Kayton, et al. (2004). "Insulinomas and expression of an insulin splice variant." Lancet 363(9406): 363-7.

Miriami, E., R. Sperling, et al. (2004). "Regulation of splicing: the importance of being translatable." RNA 10(1): 1-4.

Modrek, B. and C. J. Lee (2003). "Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss." Nat Genet 34(2): 177-80.

Moller, D. E., A. Yokota, et al. (1989). "Tissue-specific expression of two alternatively spliced insulin receptor mRNAs in man." Mol Endocrinol 3(8): 1263-9.

Moore, M. J. (2002). "RNA events. No end to nonsense." Science 298(5592): 370-1.

Mosthaf, L., K. Grako, et al. (1990). "Functionally distinct insulin receptors generated by tissue-specific alternative splicing." Embo J 9(8): 2409-13.

Mosthaf, L., B. Vogt, et al. (1991). "Altered expression of insulin receptor types A and B in the skeletal muscle of non-insulin-dependent diabetes mellitus patients." Proc Natl Acad Sci U S A 88(11): 4728-30.

Mosthaf, L., J. Eriksson, et al. (1993). "Insulin receptor isotype expression correlates with risk of non-insulin-dependent diabetes." Proc Natl Acad Sci U S A 90(7): 2633-5.

Nakai, K. and H. Sakamoto (1994). "Construction of a novel database containing aberrant splicing mutations of mammalian genes." Gene 141(2): 171-7.

Narla, G., A. Difeo, et al. (2005). "A germline DNA polymorphism enhances alternative splicing of the KLF6 tumor suppressor gene and is associated with increased prostate cancer risk." Cancer Res 65(4): 1213-22.

Nielsen, K. B., S. Sorensen, et al. (2007). "Seemingly neutral polymorphic variants may confer immunity to splicing-inactivating mutations: a synonymous SNP in exon 5 of MCAD protects from deleterious mutations in a flanking exonic splicing enhancer." Am J Hum Genet 80(3): 416-32.

Norgren, S., J. Zierath, et al. (1993). "Differences in the ratio of RNA encoding two isoforms of the insulin receptor between control and NIDDM patients. The RNA variant without Exon 11 predominates in both groups." Diabetes 42(5): 675-81.

Norgren, S., P. Arner, et al. (1994). "Insulin receptor ribonucleic acid levels and alternative splicing in human liver, muscle, and adipose tissue: tissue specificity and relation to insulin action." J Clin Endocrinol Metab 78(3): 757-62.

Norgren, S., L. S. Li, et al. (1994). "Regulation of human insulin receptor RNA splicing in HepG2 cells: effects of glucocorticoid and low glucose concentration." Biochem Biophys Res Commun 199(1): 277-84.

Novotny, I., M. Blazikova, et al. (2011). "In vivo kinetics of U4/U6.U5 tri-snRNP formation in Cajal bodies." Mol Biol Cell 22(4): 513-23.

Oberstrass, F. C., S. D. Auweter, et al. (2005). "Structure of PTB bound to RNA: specific binding and implications for splicing regulation." Science 309(5743): 2054-7.

Osheim, Y. N., O. L. Miller, Jr., et al. (1985). "RNP particles at splice junction sequences on Drosophila chorion transcripts." Cell 43(1): 143-51.

Owen, N., H. Zhou, et al. (2011). "Design principles for bifunctional targeted oligonucleotide enhancers of splicing." Nucleic Acids Res 39(16): 7194-208.

Pal, S., R. Gupta, et al. (2011). "Alternative transcription exceeds alternative splicing in generating the transcriptome diversity of cerebellar development." Genome Res 21(8): 1260-72.

Pan, Q., O. Shai, et al. (2008). "Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing." Nat Genet 40(12): 1413-5.

Pandini, G., R. Vigneri, et al. (1999). "Insulin and insulin-like growth factor-I (IGF-I) receptor overexpression in breast cancers leads to insulin/IGF-I hybrid receptor overexpression: evidence for a second mechanism of IGF-I signaling." Clin Cancer Res 5(7): 1935-44.

Paul, S., W. Dansithong, et al. (2006). "Interaction of muscleblind, CUG-BP1 and hnRNP H proteins in DM1-associated aberrant IR splicing." Embo J 25(18): 4271-83.

Perales, R. and D. Bentley (2009). ""Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions." Mol Cell 36(2): 178-91.

Pessa, H. K., C. L. Will, et al. (2008). "Minor spliceosome components are predominantly localized in the nucleus." Proc Natl Acad Sci U S A 105(25): 8655-60.

Philips, A. V., L. T. Timchenko, et al. (1998). "Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy." Science 280(5364): 737-41.

Pomeranz Krummel, D. A., C. Oubridge, et al. (2009). "Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution." Nature 458(7237): 475-80.

Pullinger, C. R., J. D. North, et al. (1989). "The apolipoprotein B gene is constitutively expressed in HepG2 cells: regulation of secretion by oleic acid, albumin, and insulin, and measurement of the mRNA half-life." J Lipid Res 30(7): 1065-77.

Puttaraju, M., S. F. Jamison, et al. (1999). "Spliceosome-mediated RNA trans-splicing as a tool for gene therapy." Nat Biotechnol 17(3): 246-52.

Raal, F. J., R. D. Santos, et al. (2010). "Mipomersen, an apolipoprotein B synthesis inhibitor, for lowering of LDL cholesterol concentrations in patients with homozygous familial hypercholesterolaemia: a randomised, double-blind, placebo-controlled trial." Lancet 375(9719): 998-1006.

Rabani, M., J. Z. Levin, et al. (2011). "Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells." Nat Biotechnol 29(5): 436-42.

Rappsilber, J., U. Ryder, et al. (2002). "Large-scale proteomic analysis of the human spliceosome." Genome Res 12(8): 1231-45.

Reed, R. and T. Maniatis (1988). "The role of the mammalian branchpoint sequence in pre-mRNA splicing." Genes Dev 2(10): 1268-76.

Robberson, B. L., G. J. Cote, et al. (1990). "Exon definition may facilitate splice site selection in RNAs with multiple exons." Mol Cell Biol 10(1): 84-94.

Roca, X., R. Sachidanandam, et al. (2003). "Intrinsic differences between authentic and cryptic 5' splice sites." Nucleic Acids Res 31(21): 6321-33.

Rodriguez-Martin, T., M. A. Garcia-Blanco, et al. (2005). "Reprogramming of tau alternative splicing by spliceosome-mediated RNA trans-splicing: implications for tauopathies." Proc Natl Acad Sci U S A 102(43): 15659-64.

Rossi, F., E. Labourier, et al. (1996). "Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I." Nature 381(6577): 80-2.

Saad, F., S. Hotte, et al. "Randomized phase II trial of Custirsen (OGX-011) in combination with docetaxel or mitoxantrone as second-line therapy in patients with metastatic castrate-resistant prostate cancer progressing after first-line docetaxel: CUOG trial P-06c." Clin Cancer Res 17(17): 5765-73.

Sanford, J. R., N. K. Gray, et al. (2004). "A novel role for shuttling SR proteins in mRNA translation." Genes Dev 18(7): 755-68.

Santoni, M. J., D. Barthels, et al. (1989). "Differential exon usage involving an unusual splicing mechanism generates at least eight types of NCAM cDNA in mouse brain." EMBO J 8(2): 385-92.

Savkur, R. S., A. V. Philips, et al. (2001). "Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy." Nat Genet 29(1): 40-7.

Savkur, R. S., A. V. Philips, et al. (2004). "Insulin receptor splicing alteration in myotonic dystrophy type 2." Am J Hum Genet 74(6): 1309-13.

Schneider, C., E. Leung, et al. (2009). "The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome." Nucleic Acids Res 37(4): 1127-40.

Schneider, M., C. L. Will, et al. (2011). "Exon definition complexes contain the tri-snRNP and can be directly converted into B-like precatalytic splicing complexes." Mol Cell 38(2): 223-35.

Sciacca, L., A. Costantino, et al. (1999). "Insulin receptor activation by IGF-II in breast cancers: evidence for a new autocrine/paracrine mechanism." Oncogene 18(15): 2471-9.

Sciacca, L., R. Mineo, et al. (2002). "In IGF-I receptor-deficient leiomyosarcoma cells autocrine IGF-II induces cell invasion and protection from apoptosis via the insulin receptor isoform A." Oncogene 21(54): 8240-50.

Seino, S. and G. I. Bell (1989). "Alternative splicing of human insulin receptor messenger RNA." Biochem Biophys Res Commun 159(1): 312-6.

Sell, S. M., D. Reese, et al. (1994). "Insulin-inducible changes in insulin receptor mRNA splice variants." J Biol Chem 269(49): 30769-72.

Sen, S., I. Talukdar, et al. (2009). "SRp20 and CUG-BP1 modulate insulin receptor exon 11 alternative splicing." Mol Cell Biol 29(3): 871-80.

Seraphin, B., L. Kretzner, et al. (1988). "A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site." EMBO J 7(8): 2533-8.

Sesti, G., M. A. Marini, et al. (1991). "Altered expression of the two naturally occurring human insulin receptor variants in isolated adipocytes of non-insulin-dependent diabetes mellitus patients." Biochem Biophys Res Commun 181(3): 1419-24.

Sesti, G., R. D'Alfonso, et al. (1995). "Peptide-based radioimmunoassay for the two isoforms of the human insulin receptor." Diabetologia 38(4): 445-53.

Sesti, G., M. Federici, et al. (2001). "Molecular mechanism of insulin resistance in type 2 diabetes mellitus: role of the insulin receptor variant forms." Diabetes Metab Res Rev 17(5): 363-73.

Shababi, M., J. Glascock, et al. (2011). "Combination of SMN trans-splicing and a neurotrophic factor increases the life span and body mass in a severe model of spinal muscular atrophy." Hum Gene Ther 22(2): 135-44.

Shabalina, S. A., A. N. Spiridonov, et al. (2010). "Connections between alternative transcription and alternative splicing in mammals." Genome Biol Evol 2: 791-9.

Shapiro, M. B. and P. Senapathy (1987). "RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression." Nucleic Acids Res 15(17): 7155-74.

Shen, H., X. Zheng, et al. (2008). "Distinct activities of the DExD/H-box splicing factor hUAP56 facilitate stepwise assembly of the spliceosome." Genes Dev 22(13): 1796-803.

Shepard, S., M. McCreary, et al. (2009). "The peculiarities of large intron splicing in animals." PLoS One 4(11): e7853.

Shin, C. and J. L. Manley (2002). "The SR protein SRp38 represses splicing in M phase cells." Cell 111(3): 407-17.

Skordis, L. A., M. G. Dunckley, et al. (2003). "Bifunctional antisense oligonucleotides provide a trans-acting splicing enhancer that stimulates SMN2 gene expression in patient fibroblasts." Proc Natl Acad Sci U S A 100(7): 4114-9.

Skruzny, M., C. Schneider, et al. (2009). "An endoribonuclease functionally linked to perinuclear mRNP quality control associates with the nuclear pore complexes." PLoS Biol 7(1): e8.

Small, E. C., S. R. Leggett, et al. (2006). "The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H box ATPase." Mol Cell 23(3): 389-99.

Smith, D. J., C. C. Query, et al. (2008). ""Nought may endure but mutability": spliceosome dynamics and the regulation of splicing." Mol Cell 30(6): 657-66.

Soares, L. M., K. Zanier, et al. (2006). "Intron removal requires proofreading of U2AF/3' splice site recognition by DEK." Science 312(5782): 1961-5.

Solis, A. S., R. Peng, et al. (2008). "Growth hormone deficiency and splicing fidelity: two serine/arginine-rich proteins, ASF/SF2 and SC35, act antagonistically." J Biol Chem 283(35): 23619-26.

Soos, M. A., J. Whittaker, et al. (1990). "Receptors for insulin and insulin-like growth factor-I can form hybrid dimers. Characterisation of hybrid receptors in transfected cells." Biochem J 270(2): 383-90.

Soret, J., M. Gabut, et al. (2006). "SR proteins as potential targets for therapy." Prog Mol Subcell Biol 44: 65-87.

Soutschek, J., A. Akinc, et al. (2004). "Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs." Nature 432(7014): 173-8.

Spies, N., C. B. Nielsen, et al. (2009). "Biased chromatin signatures around polyadenylation sites and exons." Mol Cell 36(2): 245-54.

Srirangalingam, U., B. Khoo, et al. (2010). "SDHD-related chromaffin tumours: disease localisation to genetic dysfunction." Horm Res Paediatr 73(2): 135-9.

Staffa, A., N. H. Acheson, et al. (1997). "Novel exonic elements that modulate splicing of the human fibronectin EDA exon." J Biol Chem 272(52): 33394-401.

Staley, J. P. and C. Guthrie (1998). "Mechanical devices of the spliceosome: motors, clocks, springs, and things." Cell 92(3): 315-26.

Strasser, K., S. Masuda, et al. (2002). "TREX is a conserved complex coupling transcription with messenger RNA export." Nature 417(6886): 304-8.

Sun, H. and L. A. Chasin (2000). "Multiple splicing defects in an intronic false exon." Mol Cell Biol 20(17): 6414-25.

Swanson, M. S. and G. Dreyfuss (1988). "Classification and purification of proteins of heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities." Mol Cell Biol 8(5): 2237-41.

Talerico, M. and S. M. Berget (1994). "Intron definition in splicing of small Drosophila introns." Mol Cell Biol 14(5): 3434-45.

Talmud, P., L. King-Underwood, et al. (1989). "The molecular basis of truncated forms of apolipoprotein B in a kindred with compound heterozygous hypobetalipoproteinemia." J Lipid Res 30(11): 1773-9.

Tanaka, N., A. Aronova, et al. (2007). "Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome." Genes Dev 21(18): 2312-25.

Tarn, W. Y. and J. A. Steitz (1997). "Pre-mRNA splicing: the discovery of a new spliceosome doubles the challenge." Trends Biochem Sci 22(4): 132-7.

Tarugi, P., M. Averna, et al. (2007). "Molecular diagnosis of hypobetalipoproteinemia: an ENID review." Atherosclerosis 195(2): e19-27.

Teraoka, S. N., M. Telatar, et al. (1999). "Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences." Am J Hum Genet 64(6): 1617-31.

Thrift, R. N., T. M. Forte, et al. (1986). "Characterization of lipoproteins produced by the human liver cell line, Hep G2, under defined conditions." J Lipid Res 27(3): 236-50.

Tilgner, H., C. Nikolaou, et al. (2009). "Nucleosome positioning as a determinant of exon recognition." Nat Struct Mol Biol 16(9): 996-1001.

Tsai, R. T., R. H. Fu, et al. (2005). "Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2." Genes Dev 19(24): 2991-3003.

Tseng, C. K. and S. C. Cheng (2008). "Both catalytic steps of nuclear pre-mRNA splicing are reversible." Science 320(5884): 1782-4.

Ullrich, A., J. R. Bell, et al. (1985). "Human insulin receptor and its relationship to the tyrosine kinase family of oncogenes." Nature 313(6005): 756-61.

Valadkhan, S. and J. L. Manley (2000). "A tertiary interaction detected in a human U2-U6 snRNA complex assembled in vitro resembles a genetically proven interaction in yeast." RNA 6(2): 206-19.

Valadkhan, S. and J. L. Manley (2001). "Splicing-related catalysis by protein-free snRNAs." Nature 413(6857): 701-7.

Vasudevan, S., S. W. Peltz, et al. (2002). "Non-stop decay--a new mRNA surveillance pathway." Bioessays 24(9): 785-8.

Vella, V., G. Pandini, et al. (2002). "A novel autocrine loop involving IGF-II and the insulin receptor isoform-A stimulates growth of thyroid cancer." J Clin Endocrinol Metab 87(1): 245-54.

Vester, B. and J. Wengel (2004). "LNA (locked nucleic acid): high-affinity targeting of complementary RNA and DNA." Biochemistry 43(42): 13233-41.

Vogt, B., J. M. Carrascosa, et al. (1991). "The two isotypes of the human insulin receptor (HIR-A and HIR-B) follow different internalization kinetics." Biochem Biophys Res Commun 177(3): 1013-8.

von Horn, H., M. Tally, et al. (2001). "Expression levels of insulin-like growth factor binding proteins and insulin receptor isoforms in hepatoblastomas." Cancer Lett 162(2): 253-60.

Wang, E. T., R. Sandberg, et al. (2008). "Alternative isoform regulation in human tissue transcriptomes." Nature 456(7221): 470-6.

Wang, J., J. I. Hamilton, et al. (2002). "Alternatively spliced TCR mRNA induced by disruption of reading frame." Science 297(5578): 108-10.

Wang, Z., M. E. Rolish, et al. (2004). "Systematic identification and analysis of exonic splicing silencers." Cell 119(6): 831-45.

Webby, C. J., A. Wolf, et al. (2009). "Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing." Science 325(5936): 90-3.

Webster, N. J., L. G. Evans, et al. (2004). "Assembly of splicing complexes on exon 11 of the human insulin receptor gene does not correlate with splicing efficiency in-vitro." BMC Mol Biol 5: 7.

Wesolowski, D., H. S. Tae, et al. (2011). "Basic peptide-morpholino oligomer conjugate that is very effective in killing bacteria by gene-specific and nonspecific modes." Proc Natl Acad Sci U S A 108(40): 16582-7.

West, S., N. Gromak, et al. (2004). "Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites." Nature 432(7016): 522-5.

Wheeler, T. M., K. Sobczak, et al. (2009). "Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA." Science 325(5938): 336-9.

White, M. F. (1997). "The insulin signalling system and the IRS proteins." Diabetologia 40 Suppl 2: S2-17.

Will, C. L. and R. Luhrmann (2001). "Spliceosomal UsnRNP biogenesis, structure and function." Curr Opin Cell Biol 13(3): 290-301.

Will, C. L., C. Schneider, et al. (2001). "A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site." EMBO J 20(16): 4536-46.

Wilusz, J. E., S. C. Devanney, et al. (2005). "Chimeric peptide nucleic acid compounds modulate splicing of the bcl-x gene in vitro and in vivo." Nucleic Acids Res 33(20): 6547-54.

Wu, Q. and A. R. Krainer (1999). "AT-AC pre-mRNA splicing mechanisms and conservation of minor introns in voltage-gated ion channel genes." Mol Cell Biol 19(5): 3225-36.

Xiang, S., A. Cooper-Morgan, et al. (2009). "Structure and function of the 5'-->3' exoribonuclease Rat1 and its activating partner Rai1." Nature 458(7239): 784-8.

Xiao, S. H. and J. L. Manley (1997). "Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing." Genes Dev 11(3): 334-44.

Xie, S. Y., W. Li, et al. (2011). "Correction of beta654-thalassaemia mice using direct intravenous injection of siRNA and antisense RNA vectors." Int J Hematol 93(3): 301-10.

Xu, Y. Z. and C. C. Query (2007). "Competition between the ATPase Prp5 and branch region-U2 snRNA pairing modulates the fidelity of spliceosome assembly." Mol Cell 28(5): 838-49.

Xue, Y., Y. Zhou, et al. (2009). "Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping." Mol Cell 36(6): 996-1006.

Yamaguchi, Y., J. S. Flier, et al. (1991). "Functional properties of two naturally occurring isoforms of the human insulin receptor in Chinese hamster ovary cells." Endocrinology 129(4): 2058-66.

Yamaguchi, Y., J. S. Flier, et al. (1993). "Ligand-binding properties of the two isoforms of the human insulin receptor." Endocrinology 132(3): 1132-8.

Yamanaka, S., M. E. Balestra, et al. (1995). "Apolipoprotein B mRNA-editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals." Proc Natl Acad Sci U S A 92(18): 8483-7.

Yoshimoto, R., N. Kataoka, et al. (2009). "Isolation and characterization of post-splicing lariat-intron complexes." Nucleic Acids Res 37(3): 891-902.

Yu, Y., P. A. Maroney, et al. (2008). "Dynamic regulation of alternative splicing by silencers that modulate 5' splice site competition." Cell 135(7): 1224-36.

Zamore, P. D. and M. R. Green (1989). "Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor." Proc Natl Acad Sci U S A 86(23): 9243-7.

Zhang, J., J. Cai, et al. (2007). "A genome-wide survey of alternative translational initiation events in Homo sapiens." Sci China C Life Sci 50(3): 423-8.

Zhang, X. H. and L. A. Chasin (2004). "Computational definition of sequence motifs governing constitutive exon splicing." Genes Dev 18(11): 1241-50.

Zhou, Z., L. J. Licklider, et al. (2002). "Comprehensive proteomic analysis of the human spliceosome." Nature 419(6903): 182-5.

Zhu, J., A. Mayeda, et al. (2001). "Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins." Mol Cell 8(6): 1351-61.

Zimmermann, T. S., A. C. Lee, et al. (2006). "RNAi-mediated gene silencing in non-human primates." Nature 441(7089): 111-4.

Zuo, P. and T. Maniatis (1996). "The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing." Genes Dev 10(11): 1356-68.

Chapter 6

Appendix

6.1. Antisense oligonucleotides.

Antisense Oligonucleotide	Sequence (5′- 3′)
INSR 3	mCmUmGmUmGmGmAmAmAmCmAmAmAmAmCmCmAmA
INSR 5	mCmGmCmAmCmAmGmGmUmGmAmGmUmCmAmUmAmC
INSR B	mCmGmCmCmUmUmUmGmAmGmGmAmCmAmGmAmGmG
INSR 6/7	mAmCmCmAmCmUmGmGmGmCmUmGmUmGmAmCmAmCmUmUmGmGmAmGmGmCmCmAmCmAmUmGmUmGmUmCmCmGmAmGmUmAmA
INSR 3B	mCmUmGmUmGmGmAmAmAmCmAmAmAmAmCmCmAmAmCmGmCmCmUmUmUmGmAmGmGmAmCmAmGmAmGmG
INSR 5B	mCmGmCmAmCmAmGmGmUmGmAmGmUmCmAmUmAmCmCmGmCmCmUmUmUmGmAmGmGmAmCmAmGmAmGmG
INSR 53	mCmGmCmAmCmAmGmGmUmGmAmGmUmCmAmUmAmCmCmUmGmUmGmGmAmAmAmCmAmAmAmAmCmCmAmA
INSR ex1	mGmUmGmCmCmUmGmAmAmGmAmGmGmUmUmUmUmU
INSR ex2	mCmUmAmGmGmGmUmCmCmUmCmGmGmCmAmCmCmA
INSR ex3	mUmCmGmGmCmAmCmCmAmGmUmGmCmCmUmGmAmA
INSR ex1+2	mCmUmAmGmGmGmUmCmCmUmCmGmGmCmAmCmCmAmGmUmGmCmCmUmGmAmAmGmAmGmGmUmUmUmUmU
INSR ex1+2 flu	fl-mCmUmAmGmGmGmUmCmCmUmCmGmGmCmAmCmCmA mGmUmGmCmCmUmGmAmAmGmAmGmGmUmUmUmUmU
INSR ex1+2 flu scram	fl-mGmGmGmCmUmAmUmCmGmGmCmCmUmGmUmUmAmG mGmGmUmCmGmAmCmUmGmUmCmAmCmUmAmUmCmA
INSR 53scram	mGmGmUmAmGmCmAmAmGmGmCmCmAmAmGmCmCmCmUmAmGmGmAmCmAmUmAmUmAmCmAmAmAmCmAmU
INSR ex1scram	mGmCmUmUmUmAmUmUmGmGmGmAmGmUmCmUmGmA
INSR ex2scram	mGmGmCmCmGmAmUmCmCmGmUmCmAmGmUmCmCmA
INSR ex1+2scram	mGmGmGmCmUmAmUmCmGmGmCmCmUmGmUmUmAmGmGmG mUmCmGmAmCmUmGmUmCmAmCmUmAmUmCmA
INSR TOES ex1 3′ Tail	mGmUmGmCmCmUmGmAmAmGmAmGmGmUmUmUmUmUAGGAG GACGGAGGACGGAGGACA
INSR TOES ex1 5′ Tail	AGGAGGACGGAGGACGGAGGACAmGmUmGmCmCmUmGmAmAmGmAmGmGmUmUmUmU
INSR TOES Tail only	AGGAGGACGGAGGACGGAGGACA

INSR PNA ex1	pGpTpGpCpCpTpGpApApGpApGpGpTpT
INSR PNA ex2	pGpGpGpTpCpCpTpCpGpGpCpApCpCpA
INSR PNA ex2 (18mer)	pCpTpApGpGpGpTpCpCpTpCpGpGpCpApCpCpA
INSR PNA ex4	pGpApApGpApGpGpTpTpTpTpT
INSR PNA ex5	pGpCpApCpCpApGpTpGpCpCpT
INSR PNA ex4	pCpTpApGpGpGpTpCpCpTpCpG
PNA ex1 3′ TAIL	pGpTpGpCpCpTpGpApApGpApGpGpTpT-GlyGly-ArgSerArgSer ArgSerArgSerArgSerArgSerArgSerArgSerArgSerArgSer
PNA ex1 5′ TAIL	ArgSerArgSerArgSerArgSerArgSerArgSerArgSerArgSerArgSerArgSer-GlyGly- pGpTpGpCpCpTpGpApApGpApGpGpTpT
PNA ex2 3′ TAIL	pGpGpGpUpCpCpUpCpGpGpCpApCpCpA-GlyGly-ArgSerArgSer ArgSerArgSerArgSerArgSerArgSerArgSerArgSerArgSer
PNA ex2 5′ TAIL	ArgSerArgSerArgSerArgSerArgSerArgSerArgSerArgSerArgSerArgSer-GlyGly- pGpGpGpUpCpCpUpCpGpGpCpApCpCpA
PNA Tail only	ArgSerArgSerArgSerArgSerArgSerArgSerArgSerArgSerArgSerArgSer
APOB ex1 (65-89)	mGmAmUmAmUmUmCmGmAmAmUmCmUmAmGmAmAmAmUmUm UmGmUmGmG
APOB ex2 (97-121)	mUmUmGmUmUmUmCmCmAmAmGmUmUmUmUmUmCmUmAmCm AmUmGmAmC
APOB ex3 (129-153)	mCmAmUmCmGmAmAmUmAmUmUmAmGmUmAmAmAmCmCmUm UmUmUmGmA
APOB ex4 (2319-2343)	mCmUmUmUmAmAmAmUmAmUmAmCmUmGmAmUmCmAmAmAmU mUmGmUmA
APOB ex5 (93-217) neutral	mUmGmCmUmGmUmUmUmCmUmUmUmUmUmGmGmAmGmUmCm CmAmAmAmU
APOB ex6 (2191-2215) neutral	mGmCmAmGmUmCmAmGmUmUmUmCmUmCmCmUmUmGmGmCm AmUmGmUmG

Key: mACGU = 2′-O-methyl RNA bases; ACGU = RNA bases; pACGU = peptide nucleic acid bases; fl-fluorescein; Gly = glycine; Arg = arginine; Ser = serine.

6.2 Primers.

Exon	Cycles	Primer	Forward
GAPDH	22	Forward	AAGGTGAAGGTCGGAGTCAACGG
		Reverse	CCACTTGATTTTGGAGGGATCTC
IR exon 11	32	Forward	AACCAGAGTGAGTATGAGGATTCG
		Reverse	TTCTCAAAAGGCCTGTGCTCCTCC
APOB exon 26	30	Forward	GCCATCTCGAGAGTTCCAAG
		Reverse	GTCACGGTGTGCAAATGTTC

6.3 Web-based tools

See Chapter 2.8 for information regarding the web-based tools ESE Finder, RESCUE-ESE, the PESXs server and the FAS-ESS web server.

1. **Ensembl**
   - http://www.ensembl.org/index.html
   - For gene sequences.

2. **Primer3Plus**
   - http://www.bioinformatics.nl/cgi-bin/primer3plus/primer 3plus.cgi.
   - For primer design.

3. **ESE Finder**
   - http://rulai.cshl.edu/tools/ESE2/
   - For identification of ESE sites.

4. **RESCUE-ESE**
   - http://genes.mit.edu/ burgelab/rescue-ese/index.html
   - For identification of ESE sites.

5. **PESXs server**
   - http://cubweb.biology .columbia.edu/pesx/
   - For identification of ESE/ESS sites.

5. **FAS-ESS web server**
   - http://genes.mit.edu/fas-ess/
   - For identification of ESS sites.

6. **BLAST**
   - http://blast.ncbi.nlm.nih.gov /blast/Blast.cgi

- For sequence comparison for homology.

7. Alex Dong Li's splice site finder

- http://violin.genet.sickkids.on.ca/~ali/splicesitefinder.html
- For splice site identification
- Site no longer available.

8. Reverse complement

- http://www.bioinformatics.org/sms/rev_comp.html
- For sequence reversal and/or complementation

9. siRNA Wizard v3.1

- http://www.sirnawizard.com
- For scramble sequences.

6.4 Insulin receptor Pre-mRNA sequence exon 10- exon 12

Pre-mRNA sequence showing sequence exon 10 - intron 10 - exon 11 - intron 11 - exon 12. Grey shaded areas are exonic sequences, the remaining sequence being intronic. Numbering gives nucleotide position within the IR gene.

141661 TTTCTTTCTCCAGGGCTGAAGCTGCCCTCGAGGACCTGGTCTCCACCATTCGAGTCTGAA 141720
141721 GATTCTCAGAAGCACAACCAGAGTGAGTATGAGGATTCGGCCGGCGAATGCTGCTCCTGT 141780
141781 CCAAAGACAGACTCTCAGATCCTGAAGGAGCTGGAGGAGTCCTCGTTTAGGAAGACGTTT 141840
141841 GAGGATTACCTGCACAACGTGGTTTTCGTCCCCAGGTCAGGACTTGGCGCTGGGCTCTCT 141900
141901 TAGTGGGTGCCAATTGGCTTGGTGTTGGTGGAAGGTCATTACTTAGGGACCGAGAGGTAG 141960
141961 TGGGACCCAGAGACGGCAGAAGGGTGGGTGGAGTCTGAATGGAGCCCTTTCCTGGGTGGA 142020
142021 GGAAGAGATCTTGCAGTTTCGAAATTTCGAGGGGAATTTTATCAGGAAGAGCAAATAATC 142080
142081 GCACAAAAGGGGAAAAATGGACAAAACTCAGCATTTGCCTTTCTTACCTCTCCTGAGCTG 142140
142141 GTCTCCACGTGAGTTCCAAAACAAACTCCGTTAGACGATTAACATTTTTTTTTTTCTCTC 142200
142201 TCTCTCTTTTTTTTTTTTTTTTGACAGAGTCTCACTCTATCGTCCAGGCTGGAGTGCGGT 142260
142261 GGTGCGATTTCAGCTCACCCTTTCCGGGTTCGAGTGATTCTTGTGCCTCAGCCTCCCAAG 142320
142321 TAGCTGGGACTACAGGTGCGCAGCACCACGCCCAGCTAATTTTTATATTTTTAGTAGAGA 142380
142381 CGGGGTTTTGCCATGTTGGCTAGGCTGGTCTCTAACCCCTGACCTCAGGTGATCTGCCTG 142440
142441 CCTCGGCCTCCCAGAGTGCTGGGATTATAGGTGTGAGCCACCGTGCCTGGCCGTGATTAT 142500
142501 TGAGGCTTTAGCTGAGCGCAGTGGTGTAAGCCTGTGGTCCCAGCTAGTTGGGAAGCCGAG 142560
142561 GTGGTAGGGTTGCTTGAGGCCAGCCTGGGCAACATAATGAGACCCCATCGCTACCAAAAA 142620
142621 AAAAAAAAAACAGTTTAAAAATTAGTCAGGCATGGTAATGCATGTCTGAACTCCTAGCCC 142680
142681 ATGGGGAGGCTGAGAAAGAAGGATCACAAGAGGCCAGGAGATGGAGGCTACGATAAGCGA 142740
142741 TGATGGCGCCACTGCGCTCCAGCCTGGGTGACAGAGTGAGACCCTGTCTCTAAAAAAAAA 142800
142801 TTAAAAGTATCTGTTCAGATGTTTACCAATGCTATTCATCCGAGCAGACTTTTCTCCCAG 142860
142861 AACCACCAGTGTGTAAGAATCCTCTTTTTTTTTTTTCTTTTCTTTTTCTTTTTTTGTTGC 142920
142921 ATGTTTCCCCTATGACCCTGAGTCCAAAAGAATGTTCAGGGGGACAGCCAAGGCTTCAAA 142980
142981 GTGGTTACTGTTTCACCTGTTACAAACAAAATGTTGGCTAGGCACAGTGGCTCACTCGTG 143040
143041 TAATCCCAGCACTTTGGGAGGCCGAAGTGGGAGAATCAGTTGAGGCCAAAGGTGGAGACC 143100
143101 AGCCTGAGCAACAGAGTGAGACCCTATCTCTACTAAAAAAGGAAAAAAAAAAAAAAAATA 143160
143161 GCCAGACATGTTGGTGCATGCCTTGTAGTCCCAGCTACTCCGGAGGCTGAGGCGGGAGAA 143220
143221 TTGCTGGAGCACAGGAGTTGGAGGCTGCTGTGAGCTATGATGGCACCATTGCACTCCAGC 143280
143281 CTGGGCAACAGAGCAAGACCCTGTCTCAAAAAGGAAGGAAGGGAAAAAAAGAAAGAAAGA 143340
143341 AAGGAATGGAGGGAGGGAGGGAAAGAGGAAGGAAGGAAGAGAGAAAGAAAAAGAAAGAAA 143400
143401 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAAAGAAAGAAAGGAAAGAGAG 143460
143461 AAAGAAAGGACAGAGAAAGGAAGGGGGCAAGGGAGGGAGAGAGGAGGGAAGGAAGGAAGG 143520
143521 AAAAAAGGAAAAGAAAGAAAAGAAAGAGAAAGAAAATGAAAAAAAGAAGGAAGGAAAGAG 143580
143581 AAAGAAAGCAAATGAAGGAAGGAAAGAGGGAAAAAGAGAGAAAGAAAGTAAAAGAGGAAA 143640
143641 GAAGGAAGGAAGGAGAGAGAGGGAGAGAGAAAAGAGAAAGAGAAAAGAAAAGGAGAAAGA 143700
143701 GGAAAGAAACCAGAAAGAGAGAGAGAAAGAAGAAAGAAGGGAAAGAAAAGAAAAATTCTG 143760
143761 CCAGACTTGGAGAAGTGGCTGAGTCAGTTGTGATGTCCACATGTAGTCACGTTTGACATC 143820
143821 CCAGGGCCACCTCAGCAGGCCGTCTCTGGGGAGAATTTTCTCTGATTTCTTCCCCTTCCC 143880
143881 TTGCTGGACCCCTGCACCTGCTGGGGAAGATGTAGCTCACTCCGTCTAGCAAGTGATGGG 143940
143941 AGCGAGTGGTCCAGGGTCAAAGCCAGGGTGCCCTTACTCGGACACATGTGGCCTCCAAGT 144000
144001 GTCAGAGCCCAGTGGTCTGTCTAATGAAGTTCCCTCTGTCCTCAAAGGCGTTGGTTTTGT 144060
144061 TTCCACAG**AAAAACCTCTTCAGGCACTGGTGCCGAGGACCCTAG**GTATGACTCACCTGTG 144120
144121 CGACCCCTGGTGCCTGCTCCGCGCAGGGCCGGCGGCGTGCCAGGCAGATGCCTCGGAGAA 144180
144181 CCCAGGGGTTTCTGTGGCTTTTTGCATGCGGCGGGCAGCTGTGCTGGAGAGCAGATGCTT 144240
144241 CACCAATTCAGAAATCCAATGCCTTCACTCTGAAATGAAATCTGGGCATGAATGTGGGGA 144300
144301 GAAACCTTCACTAACACACTCTTGCTAAAACATAGAATCATGGGAGTGCTTCCTGGGTAC 144360

144361 CCCCTCCCTGCCTTCTGTTTGCAGCCACTGTTTGCTCACTAAACATCTCTGCACCTCCCG 144420
144421 CGTGCATTTGCAGAGGTGGGTGGGGGTCCCCGGAGTCTGAGCTCCCCGCGGCTGGGTGCC 144480
144481 CCGACCCAGCAGCTCCTACACCATGAATGGAGGTTGATCTGGAAACAGAATATTTTCATG 144540
144541 AAAGGGCGACAGGGTATGAACAAAAGAACACCGTGTCGCTCACTGAATTCCACGGAGGAG 144600
144601 AGTCAGGGATCTCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 144660
144661 CTTCTTTCTTTTTTTCCTTCTTTCTTTCTTTCCTTCTTTCTTTTCTTTTCTTTATTCCCT 144720
144721 CCCTCTCTCCCTCCGTCCGTCCCTCCCTCCCTTCCTTCTTTCCTTCTTCTTTTTTTTTTG 144780
144781 GATGGAATTTCACTCTTGTTGCCCAGGCTAGAGTGCAGTGGTGCGATCTCGGCTCTCCAC 144840
144841 AACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATT 144900
144901 ACAGGCATGCGCCACCACGCCCGGCTAATTTTTGTACTTTTAGTAGAGACAGGGTTTCTC 144960
144961 CATGTTGGTCAGGCTGGTCTCGATCTCCCGACCTCAGGCGATCCGCCCGCCTCGGCCTCC 145020
145021 CAAAGTGCTGGAATTACAGGCGTGAGCCACCGCCCCTGGCCAGGGACGTTTCTTTTCAAC 145080
145081 TTGAGTCTTAAGAGAATGAAGCAGTATGGTCCTGGGAGAGAGAAAAAGCTTGGAGAACTT 145140
145141 GGGAGAAAGGCAAGGAGATGTTTACCTGTCGTGGCTTTTGGGAGGAACAGGAGGTGAGAC 145200
145201 TAAAGAGGCAAGTTGGGACCACTTGGAGAGATTTTCTTCCTTTAGCAGAGTGTTTATTTA 145260
145261 TTTATTTTTTCCCAGGCACCTGTGTGGCCTGATCTGCCCTCTAGTGGCCATTTGCAATAC 145320
145321 TAACTGTAGCCAAGAAGGGTCTAAAAATGTACGGGCTTGGTCCAAAGCAGATCTCAATCT 145380
145381 CAGTCCTGCTATTAATAATATTCGGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGG 145440
145441 CCGAGGCGGGCGGATCATGAGATCAGAAGATCGAGACCATCCTGGCTAACACGGTGAAAC 145500
145501 CCCGTCACTACTAAAAATACAAGAAATTAGCTGGGCGTGGTGACGCAGCTACCCCAGGCG 145560
145561 GGGCAATCACTTGAGGCCAGGAGTTCGAGACCAGCCTGACCAAGATGGTGAAACCCCGTT 145620
145621 ACTACAAGAAACATAAAAATTAGCCGGGCGTGGTGACACACGCCTGTAGTCCCAGCTACT 145680
145681 CGGGAGGCTGAGGCAGGAGAATGGCGTGAACCTGGGACGCGGAGCTTGCAGTGAGCCAAG 145740
145741 ATCGTGCCACTGCACTCCAGCCTGGGCCACAGAGCGAGACTCCGTCTCAAAAAAAATAAA 145800
145801 AATAAAATAATATTTGGGGCCACAGCTGGACTTGGTAGCTCACGCCTGTAATCCCAACAC 145860
145861 TTTGAGAGGCCCAGGCGGGGCAATCACTTGAGGCCAGGAGTTCGAGACCAGCCTGACCAA 145920
145921 GATGGTGAAACCCCGTTACTACAAGAAATATAAAAATTAGCCGGGCATGGTGACACACGC 145980
145981 CTGTAGTCCCAGCTACTTGGGAGGCCGAGGCAGGAGAATCACTTGATTTCTGGAGGCGGA 146040
146041 GGTTGCGGTGAGCCGAGATCTCGCCACTGCATTCCAGCCTGGGTGACAGAGCGAGACTCT 146100
146101 GTCTCAAAAAAAAAAAAAAAAAAGAATAAATACATAATACTAATAATAATAATAAATAAT 146160
146161 ATTTGGGGTCAGATTATTCTAGGGCACTGTCCAACAACACTGAGAAGCATCCCTGGCTTC 146220
146221 CACTCATCTGATGCCAGTAGTGTGTCTTCCCCCTAAAGTAGTAATAACCAAAAATATCTT 146280
146281 CAGACATTGCTAAATATCCCCCTGGGTGATGGGGTCAGAAGGACCCTTGGTTGAGATGAC 146340
146341 CAGGTTAGTGTGGGATTGCTGGATAATAAAATGCTGCCTAGGCCAGGTGCTGTGGCTCAT 146400
146401 GCCTGATGCCTGTAATCCCAGCACTTTGGGAGGCCCAGGCAGGCAGATCACCTGAGGTCA 146460
146461 AGAGTTTGAAACCAGCCTGGGCAACATGGCGAAACCCCGTCTCTACTAAAAATACAAAAA 146520
146521 TTAGCCGGGCATGATGATGCACGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAG 146580
146581 AATCGCTTGAACCCAGAAATTGGAAGTTACAGTGAGCCAAGATCACACCACTGCACTCCA 146640
146641 GCCTGGGCAACAGAGCAAGACTCTGTCTCAAAAAAAAAAATCAGTTGATGCTTTGTCTTT 146700
146701 TGCCACTTCATGGTCTTTTTGGTTATAATCTTTTTGGTATATAAACTCATTTGGCCTGTT 146760
146761 TGATGAGCCCATTTTACATCTCTGGTTGTTGAGATACCAATTTCTCAGTAGTTACTCATG 146820
146821 AAGCACCAACGTCTTGGTTATTTTCACGAGACCCATTTGTGCCTGTATAGAAATTATCTG 146880
146881 AATCTTGCTAATACGGACATATCTTTGGCTTGATAAGTGGTGGAGCTGTTTTCTTTTCCA 146940
146941 TTCTACAATGTTGTTAGTCTTTTAATACAAAAAGACCATAAAGAGAATCTTAAGAAATGC 147000
147001 CCCCAAAGTAGGATTTGTTTTCTGACCACACATTTTCTGATGACGCTGCAATAGAAACTA 147060
147061 AAAGTTAATTTCAAAAGATTAAACAGAAAACTTCACTTCTTGGTGGTTAAAAGTTAAACA 147120
147121 CTTACAGGCTGGGCGCAGTGACTCACACCTGTAATCGCAGCACTTTGGGAGGCCCAGGCA 147180
147181 GGCGGATCACCTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAATATGGTGAAACCCCTTC 147240
147241 TCTACTAAAAATACAAAAATTAGCCAGATGTGGTAACGTGCGCCTGTAGTCCCAGCTACT 147300
147301 CAGGAGACTGAGGCAGAAGAATCACTTGAACCTGGGAGGTGGAGGTTGCAGTGAGCCGAG 147360
147361 ATCACGCCACTGCCCTCCAGCCTGGGCAACAGAGCGAGACTCTGTCTAAAAAAAAAAAAA 147420
147421 GTTAAAAAGTTAAACACTTACATATCTATCGGGCCAAAGCAGAAATGGAAACTTCCATTG 147480
147481 AAGTCCTTTTGGAAAATAATGAAAATGTGAATATCACATACCTTAACTTGTAGAATATTG 147540
147541 TCAAAGCTTGGCTTTGATGTTACTTAGCAGACTTAAATGGATTGATAAATTAAGAAATAA 147600

147601 ACTTTAAACTATAAAACTATAAATGTGTTAGGAAGGGAAAAGCAATAACAGCTAAGCCTA 147660
147661 TGGTAAACATAATCAAAGGGAAAACACAAACCAGAAGGAAAGAAGATGACATAAAACTAG 147720
147721 CGTCTGAAGAAAATTTTAATGTGTTAAGTTCTTTGGACAATGGTGCACTTGCTAAATTTG 147780
147781 AAAGTTTCAGTAAGGTGGGAAAATTTAACTTGTCTAAATGGATTCAAAAAGAAGCAGAAA 147840
147841 AATTAAGTGGTTATGAACCTTGGAAAAAGATATGTGTCAAATAATTTCCTCAAAAAAGTC 147900
147901 CCTTGGCCCAGAGAGTTTTATAGACACGTTCTTTCAAATTGTGAGAGAGCAACTATTTCC 147960
147961 TATAATAGAAAAACTATTTTAGGGCACAGAAATCTATAAAAGTTATCCCAGTTCATTTTA 148020
148021 TGAAATGAACATAACACCCTTAAAGAACTTGACAAAGAGGTTGAGCATGGTGGCTCATGC 148080
148081 CTGTAATCCCAGTACTTTGGGAGGCTGAGGTGGGGGGCGGATCACAAGGTCAAGAGATTG 148140
148141 AGAACATTCTGGTCAACATGGTGAAACACTGTCTCTACTAACAATACAAAAATTAGCCAG 148200
148201 ACATGGTGGCACACGCCTGTAATCCCAGCTACTTGGGAGGCTGAGCCAGGAGAATCGCTT 148260
148261 GAACCTGGGAGGCAGAGGTTGCAGTGAACCGAGGTTGCGCCATTGCACACTCCAGCCTGG 148320
148321 CAACAGAGCAAAACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAGAACTTGACAAAGATA 148380
148381 CTGCACTGAACAAAAGCCACAGGCCAAGCTAATATGTGATGTTAGATTCAAATGTCCCAA 148440
148441 ATAAAACACTGCACTGGCAATATTGTTTATCTCCAAAATGCAAGGATGGTTTAAGATCAG 148500
148501 AAGATTTGATAACATCATCTAGTTCATCAAGAGGTGCAATAAGGAAGATCATATCTGTTT 148560
148561 CAATAACCTTAATTTAAAATTCCTAACTTAGAGAAAGGGGAGGAAAAAAAGAGGCTTCAT 148620
148621 AAAACTCATAATAGAATATTTCTTCTACAAGATAAATATTACCTAGGTTAAACCAACAGC 148680
148681 CAAGCTAACATCATTCTTAAAATTATTCTTAGCAGAGTTAGGAACAAAACAAGGAAGCCC 148740
148741 CATCTCATCATTACTAATTAACATTTCTCTGGGCATCCAGCCAATGTAATAAGACAAGAA 148800
148801 ACAGAAATCAGAGTTGTAAAGAATGGAATGTGGGAAATAAAGTCATTGCTATTTGCAGAG 148860
148861 CGTCTAGCTCAAGGGTAGGCAAACTATGGCCCATGGGCTAGCTGCCTGTTTTTGTAAATG 148920
148921 AACTCTTATTGGAATGCAACCCCACCCACTTGTTTACACATCATGTAGGGCTGCTTTAAG 148980
148981 CTACAATGACAGAGTAGTTGTGACAGAGGTCATTTGGCTTGCAAAACCCAAAATATTTAT 149040
149041 TATCTAGCCCTTTACGGGAAAATTTTGTTGACCCCTCTTCTAGTGTGTCATTGAAAACCC 149100
149101 AGGAAAGCCAATTATGAAACTTTTTAAATGAGTTAGAGTTTGGGAAAGTAGGTACATAGA 149160
149161 AATAAGCAAAAAGACTGTAGATACACAAGCAATAATTAGAATATAAAATAAGGGAAAAGA 149220
149221 TCCCTTTCAAAATTGCCGTAAAACCTTAGCACAACTCATAATAACCCCATCTAGAATGAC 149280
149281 TTATAAAGGAACACGTTTTAAAGTGTTACTGATATATTAAATTGTACAAATAAACATGAA 149340
149341 AGCCTGTTTCAAATCATTTAAAGAAAGCTTACTGGCAAATGGAAAAGAGTAAGTTGTATA 149400
149401 GCACGATACATAAAAGGTTAATATCCTTAATGAATAAACAACACAGGAGTAAGGAAAACA 149460
149461 CAAGAAATTCATGGAAAAACAGACAAAAGACAAGGAGGGGACAATTCTTTGGAGAATCAC 149520
149521 AGAGAACCAATAAAGGAAAAATGAAGTCACACTCAGTAATATAAAGATCCGCAAATAACG 149580
149581 TAGTCGAGTGATAGCATATTTTTAATCTCTTGGGTTAAAGATCTAGACAAACAAACATAG 149640
149641 AAAGAAAAAAAAAAAACTATTAAGGCTTAGGGGTTGCTGAGATTGTCCTGGAATGCAGTCT 149700
149701 CATGGCCTGCTGGTGCAACCACAAGTCAATATAGAGTTTCTGGACAACCATTTAGCTGTA 149760
149761 GGTATTAAGAAGCTTAAAGAGTTTCTTATCTCCCTGAACCTAGTAATATTTTCCTCTAGG 149820
149821 AATCTACTTTGGGGAAGTAATAAGACAATCAGACAATGATCAAAGGAATATTTATAACAG 149880
149881 CAAAAAAAAAAAAAAAGAGCCATGTGCGGTGGCTCATGCCTGTAATCCTGGCACTTTGGG 149940
149941 AGGCTGAGGTGGGCAGATCACCTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATGGC 150000
150001 GAAACTCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGCGGACTCCTGTAA 150060
150061 TCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATCGCATGAACTCAGGAGGCGGAGGTTGC 150120
150121 AGTGAGCCAAGATCACGCCGTCGCACTCCAGCCTGGGCGACAAGAGTGAAACTCTGTCTC 150180
150181 AAAATAAATAAATAAATAAGGAGGATGGTTAAGTGAGTGACGACTCACGTAGCAAATGGA 150240
150241 ATATTTTACAGTCATTCATCGTTTTCCAAGTTTTCTGCAGGGAACACATGCTACTTTTGT 150300
150301 AGATGATGGAGGAAAACCCTAATATTAAAAAGGAAAAGTCCCAAAATTCTCTCCTAAAAT 150360
150361 GTTATGTTGAAAGGAAAATATTTAGTAAATGCTTCCAATATAATAAGTAGACAAAGTTTA 150420
150421 CTTATAAAACAGGCTTGATCCTCATCATGTTCACACATAATATTTACAGGGAAAAAATTG 150480
150481 TCCAAAGGAAGCAGTGGGCTGTGGATTTATGAGTGATTTATATTTTCTTTTCTTTTTTTT 150540
150541 TTTTTTGAGACAGAGTCTCACTCTGTCGCCCAGGCTGGAGTGCAGTGACACGATCTCAGC 150600
150601 TCACTGCAACCTCCGCCTCCCGGGTTCAAGCAATTCTCTTGCCTCAGCCTCCTGAGTAGC 150660
150661 TGGGATTAAAGCTGTGCACCATCAGGCTCGGCTAATTTTTGTATTTTTAGTAGAGACGGG 150720
150721 GTTTCCCCATGTTGGCCAGACTGGTCTTGAACTCCTAACCTCAGGTTATCTGCCCGCCTT 150780
150781 GGCCTCCCAAAGTACTGGGATTATAGGCATGAGCCACCTCGCCCAGCCTACATTTTCTTC 150840

150841 TTTATCCTTTTTTCTTTGTATTTTAAAATTTGTTAGCTATCAACCTTTGTTGTTTAATAT 150900
150901 GTCTTCTTAAGGGATGGGGAGAAGTGTGCCAATTGCATAGATGTTTCTTGGAGTTAATCT 150960
150961 TCCTGACCTGAGTGATGTGATTCCTTTTCTAAGAATTGCAGCTTAATCACTCAAATTATA 151020
151021 CAGGTGCCCTAAGCAGCAATAAGAAGGAAAAGATGTATTTTTTCCTATTTAAACTGATGA 151080
151081 TGTCGATCAGTAGCTTGGGCCAATGAGCAAATGCCATTTGGCAGGGCTTCCTGAAAAGCT 151140
151141 AGACTCTTACGTGCTTTTGGTGTCTTATTTTTGTCACTGACCCTGGCTGGCCTATTTATC 151200
151201 TGCAAAATGAAAAACCTGAAATCGAGCATGTCGAGGGGAATATCGGTGTGTGTCTACCAT 151260
151261 TGACGGTCTTAGAAGCACTCATCCCTCACGAGGTGGGCCTGCTCTAATCCTTCAGATGCT 151320
151321 CACAAGGGTATATTGAAAGCTGCTCTGAGTGGTCATTCCTGGCAGTCTGTATTGTAATCC 151380
151381 ATGTTCCCCATTGCTGCACCCTCCTGCGCTCTGATCTTTCTTCTTAATCAAGCCTTTTAT 151440
151441 TCTCCAGTGTCACTTTTTTAAAAAAAATGATGGTGATGGTGTCATCATACATGTCCTACT 151500
151501 GTCGTTCCAG**GCCATCTCGGAAACGCAGGTCCCTTGGCGATGTTGGGAATGTGACGGTGG** 151560
151561 **CCGTGCCCACGGTGGCAGCTTTCCCCAACACTTCCTCGACCAGCGTGCCCACGAGTCCGG** 151620
151621 **AGGAGCACAGGCCTTTTGAGAAGGTGGTGAACAAGGAGTCGCTGGTCATCTCCGGCTTGC** 151680
151681 **GACACTTCACGGGCTATCGCATCGAGCTGCAGGCTTGCAACCAGGACACCCCTGAGGAAC** 151740
151741 **GGTGCAGTGTGGCAGCCTACGTCAGTGCGAGGACCATGCCTGAAG**GTAGGGCTGCTGCTC 151800

6.5 Apolipoprotein B Pre-mRNA sequence exon 25-exon 27.

Pre-mRNA sequence showing APOB sequence exon 25 - intron 25 - exon 26 - intron 26 - exon 27. Grey shaded areas are exonic sequences, the remaining sequence being intronic. Black shaded area represents 3' end of exon 26 included in the 823 base alternatively spliced product (Chapter 3.3) Numbering gives nucleotide position within the APOB gene.

31141 **CGATGGCCGGGTCAAATATACCTTGAACAAGAACAGTTTGAAAATTGAGATTCCTTTGCC** 31200
31201 **TTTTGGTGGCAAATCCTCCAGAGATCTAAAGATGTTAGAGACTGTTAGGACACCAGCCCT** 31260
31261 **CCACTTCAAGTCTGTGGGATTCCATCTGCCATCTCGAGAGTTCCAAGTCCCTACTTTTAC** 31320
31321 **CATTCCCAAGTTGTATCAACTGCAAGTGCCTCTCCTGGGTGTTCTAGACCTCTCCACGAA** 31380
31381 **TGTCTACAGCAACTTGTACAACTGGTCCGCCTCCTACAGTGGTGGCAACACCAGCACAGA** 31440
31441 **CCATTTCAGCCTTCGGGCTCGTTACCACATGAAGGCTGACTCTGTGGTTGACCTGCTTTC** 31500
31501 **CTACAATGTGCAAG**GTGAGCTATGCTCAGGTAAAGGGTGCACCGGGCTAGTTCATGGCAG 31560
31561 GCTCTAAGAGGAGAGCCTCCTCCAGGGAGGAAAGGACTTTGGCTTTCTAGCAGATAATCT 31620
31621 TCCTTGCTACTTGGAAGTCTTTTATTTTATTCAACAAATAGAAATATTTATTAAACATAT 31680
31681 CACGTGTATTAAATATTCTAGTAGGCAGTAACAGAAAGTAGACAGATAAGCCAGCAATTA 31740
31741 TAATTCAGTGTGAGAGGTGCTATGATAAAGTGTAGTATATAAGTATAAGGTAGAGTGGAA 31800
31801 GCACTCAACAAGGGAACCTAAACAAAGCCTGTGGTGGTCAGGCAAGGCTTCCTGGAGGAA 31860
31861 TGCCTTTTGCTATCAGATTTTATCTTTGCATTACAGATGGAGGAGTCTATTGCACAATTG 31920
31921 GCCCAGAAAAATGGGGCTTTATTATTGAAAGACTTTCAACATAGAGATTGCTCTGGAAAT 31980
31981 GTACTGCTTAATTTAACCAATGTCTTTTCATTTTTATGTTAG**GATCTGGAGAAACAACAT** 32040
32041 **ATGACCACAAGAATACGTTCACACTATCATATGATGGGTCTCTACGCCACAAATTTCTAG** 32100
32101 **ATTCGAATATCAAATTCAGTCATGTAGAAAAACTTGGAAACAACCCAGTCTCAAAAGGTT** 32160
32161 **TACTAATATTCGATGCATCTAGTTCCTGGGGACCACAGATGTCTGCTTCAGTTCATTTGG** 32220
32221 **ACTCCAAAAAGAAACAGCATTTGTTTGTCAAAGAAGTCAAGATTGATGGGCAGTTCAGAG** 32280
32281 **TCTCTTCGTTCTATGCTAAAGGCACATATGGCCTGTCTTGTCAGAGGGATCCTAACACTG** 32340
32341 **GCCGGCTCAATGGAGAGTCCAACCTGAGGTTTAACTCCTCCTACCTCCAAGGCACCAACC** 32400
32401 **AGATAACAGGAAGATATGAAGATGGAACCCTCTCCCTCACCTCCACCTCTGATCTGCAAA** 32460
32461 **GTGGCATCATTAAAAATACTGCTTCCCTAAAGTATGAGAACTACGAGCTGACTTTAAAAT** 32520
32521 **CTGACACCAATGGGAAGTATAAGAACTTTGCCACTTCTAACAAGATGGATATGACCTTCT** 32580
32581 **CTAAGCAAAATGCACTGCTGCGTTCTGAATATCAGGCTGATTACGAGTCATTGAGGTTCT** 32640
32641 **TCAGCCTGCTTTCTGGATCACTAAATTCCCATGGTCTTGAGTTAAATGCTGACATCTTAG** 32700
32701 **GCACTGACAAAATTAATAGTGGTGCTCACAAGGCGACACTAAGGATTGGCCAAGATGGAA** 32760
32761 **TATCTACCAGTGCAACGACCAACTTGAAGTGTAGTCTCCTGGTGCTGGAGAATGAGCTGA** 32820
32821 **ATGCAGAGCTTGGCCTCTCTGGGGCATCTATGAAATTAACAACAAATGGCCGCTTCAGGG** 32880
32881 **AACACAATGCAAAATTCAGTCTGGATGGGAAAGCCGCCCTCACAGAGCTATCACTGGGAA** 32940
32941 **GTGCTTATCAGGCCATGATTCTGGGTGTCGACAGCAAAAACATTTTCAACTTCAAGGTCA** 33000
33001 **GTCAAGAAGGACTTAAGCTCTCAAATGACATGATGGGCTCATATGCTGAAATGAAATTTG** 33060
33061 **ACCACACAAACAGTCTGAACATTGCAGGCTTATCACTGGACTTCTCTTCAAAACTTGACA** 33120
33121 **ACATTTACAGCTCTGACAAGTTTTATAAGCAAACTGTTAATTTACAGCTACAGCCCTATT** 33180
33181 **CTCTGGTAACTACTTTAAACAGTGACCTGAAATACAATGCTCTGGATCTCACCAACAATG** 33240
33241 **GGAAACTACGGCTAGAACCCCTGAAGCTGCATGTGGCTGGTAACCTAAAAGGAGCCTACC** 33300
33301 **AAAATAATGAAATAAAACACATCTATGCCATCTCTTCTGCTGCCTTATCAGCAAGCTATA** 33360
33361 **AAGCAGACACTGTTGCTAAGGTTCAGGGTGTGGAGTTTAGCCATCGGCTCAACACAGACA** 33420
33421 **TCGCTGGGCTGGCTTCAGCCATTGACATGAGCACAAACTATAATTCAGACTCACTGCATT** 33480
33481 **TCAGCAATGTCTTCCGTTCTGTAATGGCCCCGTTTACCATGACCATCGATGCACATACAA** 33540
33541 **ATGGCAATGGGAAACTCGCTCTCTGGGGAGAACATACTGGGCAGCTGTATAGCAAATTCC** 33600
33601 **TGTTGAAAGCAGAACCTCTGGCATTTACTTTCTCTCATGATTACAAAGGCTCCACAAGTC** 33660

33661 ATCATCTCGTGTCTAGGAAAAGCATCAGTGCAGCTCTTGAACACAAAGTCAGTGCCCTGC 33720
33721 TTACTCCAGCTGAGCAGACAGGCACCTGGAAACTCAAGACCCAATTTAACAACAATGAAT 33780
33781 ACAGCCAGGACTTGGATGCTTACAACACTAAAGATAAAATTGGCGTGGAGCTTACTGGAC 33840
33841 GAACTCTGGCTGACCTAACTCTACTAGACTCCCCAATTAAAGTGCCACTTTTACTCAGTG 33900
33901 AGCCCATCAATATCATTGATGCTTTAGAGATGAGAGATGCCGTTGAGAAGCCCCAAGAAT 33960
33961 TTACAATTGTTGCTTTTGTAAAGTATGATAAAAACCAAGATGTTCACTCCATTAACCTCC 34020
34021 CATTTTTTGAGACCTTGCAAGAATATTTTGAGAGGAATCGACAAACCATTATAGTTGTAC 34080
34081 TGGAAAACGTACAGAGAAACCTGAAGCACATCAATATTGATCAATTTGTAAGAAAATACA 34140
34141 GAGCAGCCCTGGGAAAACTCCCACAGCAAGCTAATGATTATCTGAATTCATTCAATTGGG 34200
34201 AGAGACAAGTTTCACATGCCAAGGAGAAACTGACTGCTCTCACAAAAAAGTATAGAATTA 34260
34261 CAGAAAATGATATACAAATTGCATTAGATGATGCCAAAATCAACTTTAATGAAAAACTAT 34320
34321 CTCAACTGCAGACATATATGATACAATTTGATCAGTATATTAAAGATAGTTATGATTTAC 34380
34381 ATGATTTGAAAATAGCTATTGCTAATATTATTGATGAAATCATTGAAAAATTAAAAAGTC 34440
34441 TTGATGAGCACTATCATATCCGTGTAAATTTAGTAAAAACAATCCATGATCTACATTTGT 34500
34501 TTATTGAAAATATTGATTTTAACAAAAGTGGAAGTAGTACTGCATCCTGGATTCAAAATG 34560
34561 TGGATACTAAGTACCAAATCAGAATCCAGATACAAGAAAAACTGCAGCAGCTTAAGAGAC 34620
34621 ACATACAGAATATAGACATCCAGCACCTAGCTGGAAAGTTAAAACAACACATTGAGGCTA 34680
34681 TTGATGTTAGAGTGCTTTTAGATCAATTGGGAACTACAATTTCATTTGAAAGAATAAATG 34740
34741 ACATTCTTGAGCATGTCAAACACTTTGTTATAAATCTTATTGGGGATTTTGAAGTAGCTG 34800
34801 AGAAAATCAATGCCTTCAGAGCCAAAGTCCATGAGTTAATCGAGAGGTATGAAGTAGACC 34860
34861 AACAAATCCAGGTTTTAATGGATAAATTAGTAGAGTTGGCCCACCAATACAAGTTGAAGG 34920
34921 AGACTATTCAGAAGCTAAGCAATGTCCTACAACAAGTTAAGATAAAAGATTACTTTGAGA 34980
34981 AATTGGTTGGATTTATTGATGATGCTGTCAAGAAGCTTAATGAATTATCTTTTAAAACAT 35040
35041 TCATTGAAGATGTTAACAAATTCCTTGACATGTTGATAAAGAAATTAAAGTCATTTGATT 35100
35101 ACCACCAGTTTGTAGATGAAACCAATGACAAAATCCGTGAGGTGACTCAGAGACTCAATG 35160
35161 GTGAAATTCAGGCTCTGGAACTACCACAAAAAGCTGAAGCATTAAAACTGTTTTTAGAGG 35220
35221 AAACCAAGGCCACAGTTGCAGTGTATCTGGAAAGCCTACAGGACACCAAAATAACCTTAA 35280
35281 TCATCAATTGGTTACAGGAGGCTTTAAGTTCAGCATCTTTGGCTCACATGAAGGCCAAAT 35340
35341 TCCGAGAGACCCTAGAAGATACACGAGACCGAATGTATCAAATGGACATTCAGCAGGAAC 35400
35401 TTCAACGATACCTGTCTCTGGTAGGCCAGGTTTATAGCACACTTGTCACCTACATTTCTG 35460
35461 ATTGGTGGACTCTTGCTGCTAAGAACCTTACTGACTTTGCAGAGCAATATTCTATCCAAG 35520
35521 ATTGGGCTAAACGTATGAAAGCATTGGTAGAGCAAGGGTTCACTGTTCCTGAAATCAAGA 35580
35581 CCATCCTTGGGACCATGCCTGCCTTTGAAGTCAGTCTTCAGGCTCTTCAGAAAGCTACCT 35640
35641 TCCAGACACCTGATTTTATAGTCCCCCTAACAGATTTGAGGATTCCATCAGTTCAGATAA 35700
35701 ACTTCAAAGACTTAAAAAATATAAAAATCCCATCCAGGTTTTCCACACCAGAATTTACCA 35760
35761 TCCTTAACACCTTCCACATTCCTTCCTTTACAATTGACTTTGTAGAAATGAAAGTAAAGA 35820
35821 TCATCAGAACCATTGACCAGATGCTGAACAGTGAGCTGCAGTGGCCCGTTCCAGATATAT 35880
35881 ATCTCAGGGATCTGAAGGTGGAGGACATTCCTCTAGCGAGAATCACCCTGCCAGACTTCC 35940
35941 GTTTACCAGAAATCGCAATTCCAGAATTCATAATCCCAACTCTCAACCTTAATGATTTTC 36000
36001 AAGTTCCTGACCTTCACATACCAGAATTCCAGCTTCCCCACATCTCACACACAATTGAAG 36060
36061 TACCTACTTTTGGCAAGCTATACAGTATTCTGAAAATCCAATCTCCTCTTTTCACATTAG 36120
36121 ATGCAAATGCTGACATAGGGAATGGAACCACCTCAGCAAACGAAGCAGGTATCGCAGCTT 36180
36181 CCATCACTGCCAAAGGAGAGTCCAAATTAGAAGTTCTCAATTTTGATTTTCAAGCAAATG 36240
36241 CACAACTCTCAAACCCTAAGATTAATCCGCTGGCTCTGAAGGAGTCAGTGAAGTTCTCCA 36300
36301 GCAAGTACCTGAGAACGGAGCATGGGAGTGAAATGCTGTTTTTTGGAAATGCTATTGAGG 36360
36361 GAAAATCAAACACAGTGGCAAGTTTACACACAGAAAAAAATACACTGGAGCTTAGTAATG 36420
36421 GAGTGATTGTCAAGATAAACAATCAGCTTACCCTGGATAGCAACACTAAATACTTCCACA 36480
36481 AATTGAACATCCCCAAACTGGACTTCTCTAGTCAGGCTGACCTGCGCAACGAGATCAAGA 36540
36541 CACTGTTGAAAGCTGGCCACATAGCATGGACTTCTTCTGGAAAAGGGTCATGGAAATGGG 36600
36601 CCTGCCCCAGATTCTCAGATGAGGGAACACATGAATCACAAATTAGTTTCACCATAGAAG 36660
36661 GACCCCTCACTTCCTTTGGACTGTCCAATAAGATCAATAGCAAACACCTAAGAGTAAACC 36720
36721 AAAACTTGGTTTATGAATCTGGCTCCCTCAACTTTTCTAAACTTGAAATTCAATCACAAG 36780
36781 TCGATTCCCAGCATGTGGGCCACAGTGTTCTAACTGCTAAAGGCATGGCACTGTTTGGAG 36840
36841 AAGGGAAGGCAGAGTTTACTGGGAGGCATGATGCTCATTTAAATGGAAAGGTTATTGGAA 36900

36901 CTTTGAAAAATTCTCTTTTCTTTTCAGCCCAGCCATTTGAGATCACGGCATCCACAAACA 36960
36961 ATGAAGGGAATTTGAAAGTTCGTTTTCCATTAAGGTTAACAGGGAAGATAGACTTCCTGA 37020
37021 ATAACTATGCACTGTTTCTGAGTCCCAGTGCCCAGCAAGCAAGTTGGCAAGTAAGTGCTA 37080
37081 GGTTCAATCAGTATAAGTACAACCAAAATTTCTCTGCTGGAAACAACGAGAACATTATGG 37140
37141 AGGCCCATGTAGGAATAAATGGAGAAGCAAATCTGGATTTCTTAAACATTCCTTTAACAA 37200
37201 TTCCTGAAATGCGTCTACCTTACACAATAATCACAACTCCTCCACTGAAAGATTTCTCTC 37260
37261 TATGGGAAAAAACAGGCTTGAAGGAATTCTTGAAAACGACAAAGCAATCATTTGATTTAA 37320
37321 GTGTAAAAGCTCAGTATAAGAAAAACAAACACAGGCATTCCATCACAAATCCTTTGGCTG 37380
37381 TGCTTTGTGAGTTTATCAGTCAGAGCATCAAATCCTTTGACAGGCATTTTGAAAAAAACA 37440
37441 GAAACAATGCATTAGATTTTGTCACCAAATCCTATAATGAAACAAAAATTAAGTTTGATA 37500
37501 AGTACAAAGCTGAAAAATCTCACGACGAGCTCCCCAGGACCTTTCAAATTCCTGGATACA 37560
37561 CTGTTCCAGTTGTCAATGTTGAAGTGTCTCCATTCACCATAGAGATGTCGGCATTCGGCT 37620
37621 ATGTGTTCCCAAAAGCAGTCAGCATGCCTAGTTTCTCCATCCTAGGTTCTGACGTCCGTG 37680
37681 TGCCTTCATACACATTAATCCTGCCATCATTAGAGCTGCCAGTCCTTCATGTCCCTAGAA 37740
37741 ATCTCAAGCTTTCTCTTCCAGATTTCAAGGAATTGTGTACCATAAGCCATATTTTTATTC 37800
37801 CTGCCATGGGCAATATTACCTATGATTTCTCCTTTAAATCAAGTGTCATCACACTGAATA 37860
37861 CCAATGCTGAACTTTTTAACCAGTCAGATATTGTTGCTCATCTCCTTTCTTCATCTTCAT 37920
37921 CTGTCATTGATGCACTGCAGTACAAATTAGAGGGCACCACAAGATTGACAAGAAAAAGGG 37980
37981 GATTGAAGTTAGCCACAGCTCTGTCTCTGAGCAACAAATTTGTGGAGGGTAGTCATAACA 38040
38041 GTACTGTGAGCTTAACCACGAAAAATATGGAAGTGTCAGTGGCAACAACCACAAAAGCCC 38100
38101 AAATTCCAATTTTGAGAATGAATTTCAAGCAAGAACTTAATGGAAATACCAAGTCAAAAC 38160
38161 CTACTGTCTCTTCCTCCATGGAATTTAAGTATGATTTCAATTCTTCAATGCTGTACTCTA 38220
38221 CCGCTAAAGGAGCAGTTGACCACAAGCTTAGCTTGGAAAGCCTCACCTCTTACTTTTCCA 38280
38281 TTGAGTCATCTACCAAAGGAGATGTCAAGGGTTCGGTTCTTTCTCGGGAATATTCAGGAA 38340
38341 CTATTGCTAGTGAGGCCAACACTTACTTGAATTCCAAGAGCACACGGTCTTCAGTGAAGC 38400
38401 TGCAGGGCACTTCCAAAATTGATGATATCTGGAACCTTGAAGTAAAAGAAAATTTTGCTG 38460
38461 GAGAAGCCACACTCCAACGCATATATTCCCTCTGGGAGCACAGTACGAAAAACCACTTAC 38520
38521 AGCTAGAGGGCCTCTTTTTCACCAACGGAGAACATACAAGCAAAGCCACCCTGGAACTCT 38580
38581 CTCCATGGCAAATGTCAGCTCTTGTTCAGGTCCATGCAAGTCAGCCCAGTTCCTTCCATG 38640
38641 ATTTCCCTGACCTTGGCCAGGAAGTGGCCCTGAATGCTAACACTAAGAACCAGAAGATCA 38700
38701 GATGGAAAAATGAAGTCCGGATTCATTCTGGGTCTTTCCAGAGCCAGGTCGAGCTTTCCA 38760
38761 ATGACCAAGAAAAGGCACACCTTGACATTGCAGGATCCTTAGAAGGACACCTAAGGTTCC 38820
38821 TCAAAAATATCATCCTACCAGTCTATGACAAGAGCTTATGGGATTTCCTAAAGCTGGATG 38880
38881 TAACCACCAGCATTGGTAGGAGACAGCATCTTCGTGTTTCAACTGCCTTTGTGTACACCA 38940
38941 AAAACCCCAATGGCTATTCATTCTCCATCCCTGTAAAAGTTTTGGCTGATAAATTCATTA 39000
39001 TTCCTGGGCTGAAACTAAATGATCTAAATTCAGTTCTTGTCATGCCTACGTTCCATGTCC 39060
39061 CATTTACAGATCTTCAGGTTCCATCGTGCAAACTTGACTTCAGAGAAATACAAATCTATA 39120
39121 AGAAGCTGAGAACTTCATCATTTGCCCTCAACCTACCAACACTCCCCGAGGTAAAATTCC 39180
39181 CTGAAGTTGATGTGTTAACAAAATATTCTCAACCAGAAGACTCCTTGATTCCCTTTTTTG 39240
39241 AGATAACCGTGCCTGAATCTCAGTTAACTGTGTCCCAGTTCACGCTTCCAAAAAGTGTTT 39300
39301 CAGATGGCATTGCTGCTTTGGATCTAAATGCAGTAGCCAACAAGATCGCAGACTTTGAGT 39360
39361 TGCCCACCATCATCGTGCCTGAGCAGACCATTGAGATTCCCTCCATTAAGTTCTCTGTAC 39420
39421 CTGCTGGAATTGTCATTCCTTCCTTTCAAGCACTGACTGCACGCTTTGAGGTAGACTCTC 39480
39481 CCGTGTATAATGCCACTTGGAGTGCCAGTTTGAAAAACAAAGCAGATTATGTTGAAACAG 39540
39541 TCCTGGATTCCACATGCAGCTCAACCGTACAGTTCCTAGAATATGAACTAAATGGTAAGA 39600
39601 AATATCCTGCCTCCTCTCCTAGATACTGTATATTTTCAATGAGAGTTATGAGTAAATAAT 39660
39661 TATGTATTTAGTTGTGAGTAGATGTACAATTACTCAATGTCACAAAATTTTAAGTAAGAA 39720
39721 AAGAGATACATGTATACCCTACACGTAAAAACCAAACTGTAGAAAATCTAGTGTCATTCA 39780
39781 AGACAAACAGCTTTAAAGAAAATGGATTTTTCTGTAATTATTTTAGGACTAACAATGTCT 39840
39841 TTTAACTATTTATTTTAAAATAAGTGTGAGCTGTACATTGCATATTTTAAACACAAGTGA 39900
39901 AATATCTGGTTAGGATAGAATTCTCCCAGTTTTCACAATGAAAACATCAACGTCCTACTG 39960
39961 TTATGAATCTAATAAAATACAAAATCTCTCCTATACAGTTTTGGGAACACACAAAATCGA 40020
40021 AGATGGTACGTTAGCCTCTAAGACTAAAGGAACATTTGCACACCGTGACTTCAGTGCAGA 40080
40081 ATATGAAGAAGATGGCAAATATGAAGGACTTCAGTATGGAGCTTTTATTGAATTGAAACC 40140