Abstract

RNA-binding proteins play a central role in the post-transcriptional regulation of gene expression; however, little is known about the endogenous transcripts to which they bind. Here, I have used the ultra-violet cross-linking and immuno-precipitation (CLIP) technique to identify RNA targets directly bound to two RNA-binding proteins: Acinus and hnRNP A1.

Acinus (apoptotic chromatin condensation inducer in the nucleus) contains a region that is homologous to the RNA binding domain of the Drosophila splicing regulator sex-lethal, and a serine and arginine rich region similar to that seen in the SR family of proteins, which function extensively in splicing. Furthermore it is a component of the multi-protein spliceosome complex, and I have demonstrated it can directly bind polyadenylated RNA. I have shown that Acinus displays a diffuse nuclear localisation pattern, however, overexpression of an epitope-tagged protein results in its accumulation in enlarged nuclear speckles. Together these results suggest a role in pre-mRNA splicing.

Acinus is cleaved during apoptosis by caspase-3, resulting in a truncated protein with chromatin condensation inducing activity (Sahara *et al.*, 1999). Accordingly, I have demonstrated that overexpression of epitope-tagged Acinus results in an increased number of cells exhibiting an apoptotic phenotype. The proteolytic fragment contains the RNA binding region, and to determine if the role of Acinus in apoptosis is mediated by RNA interactions I utilised CLIP to identify *in vivo* RNA targets.

I have identified several mRNA targets of Acinus and found that the binding sites in those mRNA targets predominantly map to constitutively expressed exons. This is in agreement with the exon junction complex, of which Acinus is a component, being deposited on mRNAs after splicing. These results may indicate that Acinus is a core RNA binding factor of the exon junction complex.

To complement this approach, I also performed CLIP with a known alternative splicing regulator, hnRNP A1. In this manner, the binding site preferences could be compared between the two proteins. As expected, the majority of hnRNP A1 binding

sites are located in introns, corresponding with their identified role of antagonizing pre-mRNA splicing by binding intronic splicing elements. Interestingly, a number of the CLIP tags are located in, or adjacent to, alternatively spliced events suggesting a role for hnRNP A1 in the regulation of alternative splicing of these specific pre-mRNAs. In addition to pre-mRNA splicing hnRNP A1 also functions in the cellular stress response. Upon environmental stresses it relocates to the cytoplasm and accumulates in cytoplasmic foci known as stress granules (Guil *et al.*, 2006). Here I show some of the targets identified by CLIP are regulated by hnRNP A1 in times of cellular stress.

In summary, I have identified two novel subsets of RNAs, bound by Acinus or hnRNP A1 *in vivo*. I have shown these proteins exhibit distinct binding preferences, which correspond to their biological function. This work is consistent with hnRNP A1 acting as an alternative splicing regulator, and provides evidence for a dual role of Acinus in mRNA splicing and apoptosis. This study also demonstrates the power of the CLIP technique, as identification of *in vivo* RNA targets allows greater understanding of the mechanisms by which RNA-binding proteins exert their regulatory control.

Declaration

I declare that the work presented in this PhD is my own and has not been submitted for any other degree, unless otherwise stated. Parts of the introduction are taken from the Biochemical Journal Review Article (Long and Caceres, in press), of which I am the author.

Jennifer Long

Acknowledgements

I thank all members of the Caceres lab, past and present, for their support and advice, and for extravaganzas to Krakow and Barcelona!
Thanks to Javier for the opportunity (sorry, had to get that one in...).

Thanks to Pedro Coutinho for bioinformatic analyses of CLIP tags, and thanks to Colin Semple for the generation of a transcriptome map and introducing me to the world of Perl programming.
Thanks to Craig for teaching me the merits of Adobe Illustrator and InDesign.
Thanks to Stewart and his team for technical support, including the thousands of sequencing reactions which have made this PhD.

I thank all members of CGE over the last four years who have helped me by various means; lending me reagents, giving technical advice, introducing new music, feeding me and listening to my continuous PhD whinging...!

Ahoy to Bloody Morgan Vane and Pirate Riggers! The pirate duo who made life in the unit a little more interesting, you never could tell if an Arrghh! was around the corner...

The student crowd – Leeanne, Jon, Catherine, Joel, Andy, Rachel, Marion, Hannah, Ross, Jamie, Luke, G-Lover, Duncan, amongst others – have always been there to put the world to rights over a drink or two, or have a laugh at the expense of our glorious leaders in the student panto. And who can forget the one and only Midori Club?! Rock on David, Sally and Catherine!

A big thanks to members of the Hen Party Bay – Shelagh, Lizbeth, Diana, Catherine, Heidi, and honourary member Lorna - for some interesting chat, relieving the science tedium, indulging the different projectiles launched at them and the occasional cocktail. A special thanks to Shelagh for everything; lunchtime adventures, coffee

and scones, sandcastles, teaching me RNA FISH and sucrose gradients, putting up with my endless questions and tantrums, and the never-ending quest for the perfect chocolate brownie.

And a thank you to everyone else I've not named but have helped me or listened to my ramblings over the last four years.

Finally, I thank my parents for their continuous support and encouragement.

Abbreviations

aa	amino acids
Acinus	Apoptotic chromatin condensation inducer in the nucleus
AIF	Apoptosis Inducing Factor
ASAP	Apoptosis and Splicing Associated Protein
AS-NMD	Alternative Splicing-NMD
BH	Bcl-2 Homology
bp	base pairs
BP	Branch-Point
CAD	Caspase-Activated DNase
CARD	Caspase Recruitment Domain
CDK	Cyclin-Dependent Kinase
ChIP	Chromatin Immuno-Precipitation
CLIP	Cross-Linking and Immuno-Precipitation
CPSF	Cleavage and Polyadenylation Specificity Factor
CstF	Cleavage stimulation Factor
CTD	Carboxyl-Terminal Domain
CTE	Constitutive Transport Element
cTNT	Cardiac Troponin T
^{o}C	degrees Celsius
DAPI	4', 6-diamidino-2-phenylindole
DD	Death Domain
DED	Death Effector Domain
dH_2O	distilled water
DNA	Deoxyribonucleic Acid
EDTA	Ethylene diamine tetra-acetic acid
eIF4E	Eukaryotic Initiation Factor 4E

EJC	Exon Junction Complex
ESE	Exonic Splicing Enhancer
ESS	Exonic Splicing Silencer
EST	Expressed Sequence Tag
FCS	Foetal Calf Serum
FISH	Fluorescent *in situ* Hybridisation
FRAP	Fluorescence Recovery After Photo-bleaching
FRET	Fluorescence Resonance Energy Transfer
G	G centrifugal force
g	gram
H_2O	water
HITS-CLIP	High Throughput Sequencing – CLIP
HnRNP	Heterogeneous nuclear ribonucleoparticle
IP	Immuno-Precipitation
IRES	Internal Ribosome Entry Site
ISE	Intronic Splicing Enhancer
ISS	Intronic Splicing Silencer
kb	kilo base
kDa	kiloDalton
L	Litre
LB	Luria Broth
M	Molar
MAP	Mitogen activated protein
ml	millilitre
mM	millimolar

µl	microlitre
mRNA	messenger RNA
mTOR	mammalian Target Of Rapamycin
MTS	3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfopheny)-2H-tetrazolium
ncRNA	non-coding RNA
NLS	Nuclear Localisation Signal
nm	nanometre
NMD	Nonsense Mediated Decay
nt	nucleotide
ORF	Open Reading Frame
OSM	Osmotic shock
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PESE	Putative Exonic Splicing Enhancer
PKC	Protein Kinase C
PTC	Premature Termination Codon
P-TEFb	Positive Transcription Elongation Factor b
RESCUE	Relative Enhancer and Silencer Classification by Unanimous Enrichment
RIN	RNA Integrity Number
RIP	Ribonucleoprotein Immuno-precipitation
RNA	Ribonucleic Acid
RNAi	RNA interference
RNAP II	RNA polymerase II
RNase	Ribonuclease
rpm	revolutions per minute
RRM	RNA Recognition Motif

rRNA	ribosomal RNA
RT-PCR	Reverse Transcriptase – PCR
SC35	Spliceosomal Component 35
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	SDS – PolyAcrylamide Gel Electrophoresis
SELEX	Selected evolution of ligands through exponential enrichment
SF2/ASF	Splicing Factor 2/ Alternative Splicing Factor
SG	Stress Granule
shRNAi	shorthairpin RNA interference
SMA	Spinal Muscular Atrophy
snoRNA	small nucleolar RNA
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
ss	splice site
STS	Staurosporine
Sxl	Sex lethal
TCA	Trichloroacetic acid
Tra	Transformer
Tris	Tris(hydroxymethyl)-amino-methane
Trn	Transportin
tRNA	transfer RNA
U2AF	U2 snRNP auxiliary factor
UTR	Untranslated Region
UV	Ultraviolet

Single Letter Amino Acid Code

A	**Alanine**	**Aliphatic**	**L**	**Leucine**	**Aliphatic**
R	**Arginine**	**Basic**	**K**	**Lysine**	**Basic**
N	**Aspargine**	**Acidic**	**M**	**Methionine**	**Sulphur containing**
D	**Aspartate**	**Acidic**	**F**	**Phenylalanine**	**Aromatic**
C	**Cysteine**	**Sulphur containing**	**P**	**Proline**	**Imino**
Q	**Glutamine**	**Acidic**	**S**	**Serine**	**Non-Aromatic**
E	**Glutamate**	**Acidic**	**T**	**Threonine**	**Non-Aromatic**
G	**Glycine**	**Aliphatic**	**W**	**Tryptophan**	**Aromatic**
H	**Histidine**	**Basic**	**Y**	**Tyrosine**	**Aromatic**
I	**Isoleucine**	**Aliphatic**	**V**	**Valine**	**Aliphatic**

Table of contents

Abstract ..... **i**
Declaration ..... **iii**
Acknowledgements ..... **iv**
Abbreviations ..... **vi**
Table of contents ..... **xi**
List of Figures ..... **xiv**
List of Tables ..... **xvi**

CHAPTER 1: INTRODUCTION **1**
1.1. Gene expression and RNA processing **2**
1.1. Pre-mRNA splicing **3**
1.2. The spliceosome **4**
1.3. The Exon Junction Complex **7**
1.4. Alternative splicing **8**
1.5. The SR protein family **11**
1.6. Co-transcriptional splicing **15**
1.7. The heterogeneous nuclear ribonucleoparticle (hnRNP) family **16**
1.8. Roles of classical SR proteins in constitutive and alternative splicing **17**
1.9. Identification of RNA-binding protein targets **21**
1.10. SR proteins have functional specificity **23**
1.11. Post-splicing activities of SR proteins **25**
1.12. Splicing regulation by phosphorylation **27**
1.13. Splicing and human disease **29**
1.14. Splicing and apoptosis **30**
1.15. Acinus **33**
1.16. Project outline **38**

CHAPTER 2: MATERIALS AND METHODS **40**
2.1. Materials **41**
2.2. General Molecular Biology Methods **42**
Gel electrophoresis 42
Phenol/chloroform extraction and ethanol precipitation 43
PCR purification by gel extraction 43
Nucleic acid quantification 43
Sequencing 43
RNA isolation 44
pGEM-T easy cloning 44
Bacterial transformations 44
Bacterial Cell culture 44
Isolation of plasmid DNA 45
2.3. Cell Culture Methods **45**
Cell Culture 45
Cell counting 45

DNA Transfection 45
RNA Transfection 46
Induction of apoptosis 46
Cell viability assay 46
2.4. Experimental procedures 47
Plasmid constructs 47
Immunofluorescence Microscopy 47
Cell Fractionation 48
In vivo UV Cross-Linking followed by poly A^+ affinity selection 48
Immuno-Precipitation (IP) 49
Western Blot Analysis 49
Cross Linking and Immuno-Precipitation (CLIP) assay 50
Primer design 51
Statistical tests 53
2.5. Bioinformatics 53
Characterisation of the CLIP tags 53
Identification of the hnRNP A1 binding site 53
Modelling CLIP in silico 53

CHAPTER 3: ACINUS CHARACTERISATION 55
3.1. Introduction 56
3.2. Results 57
Acinus is a nuclear protein 57
Expression analysis of Acinus 62
Acinus is an RNA-binding protein 63
Classification of apoptotic stages 65
RNAi mediated knockdown of Acinus 66
Over-expression of Acinus causes apoptosis 68
3.3. Discussion 72

CHAPTER 4: ACINUS CLIP 76
4.1. Introduction 77
4.2. Cross-linking and immuno-precipitation (CLIP) technique 77
4.3. Results 79
Optimisation of CLIP for Acinus 79
Identification of Acinus targets 82
Analysis of sites to which Acinus binds 88
Acinus is an alternative splicing regulator 91
4.4. Discussion 99

CHAPTER 5: HNRNP A1 CLIP 102
5.1. Introduction 103
5.2. Results 104
HnRNP A1 CLIP 104
Identification of hnRNP A1 targets 106
Analysis of sites to which hnRNP A1 binds 112

HnRNP A1 is involved in the stress response 116
HnRNP A1 modulates the alternative splicing of several transcripts 117
RNA binding proteins show different binding profiles with respect to gene structure 129
5.3. Discussion 131

CHAPTER 6: DISCUSSION AND FUTURE WORK 137

REFERENCES 146

APPENDICES 169

List of Figures

Figure 1.1. Splice site elements. 3
Figure 1.2. Splicing takes place in two transesterification reactions. 4
Figure 1.3. Spliceosome assembly. 5
Figure 1.4. Protein components of the exon junction complex (EJC). 8
Figure 1.5. Elementary alternative splicing events. 10
Figure 1.6. Schematic diagram of SR and SR-related proteins. 12
Figure 1.7. Features of hnRNP A1. 17
Figure 1.8. Roles of SR proteins in splice site selection. 19
Figure 1.9. Role of SF2/ASF in translation. 26
Figure 1.10. Death Receptor Signalling. 31
Figure 1.11. ACIN1 is located on chromosome 14 and encodes three isoforms of Acinus generated by use of alternate promoters and alternative splicing. 35
Figure 1.12. Schematic of Acinus isoforms. 36
Figure 1.13. Acinus is conserved across metazoa. 37
Figure 3.1. Morphological changes associated with apoptosis. 56
Figure 3.2. Acinus is localised to the nucleus. 58
Figure 3.3. Acinus localises to nuclear speckles. 59
Figure 3.4. Acinus displays a nuclear localisation when over-expressed. 61
Figure 3.5. Acinus is widely expressed. 62
Figure 3.6. Acinus binds RNA. 64
Figure 3.7. Cells treated with staurosporine (STS) undergo apoptosis. 65
Figure 3.8. Classification of apoptotic stages. 66
Figure 3.9. Design and testing of Acinus siRNAs. 67
Figure 3.10. Overexpression of Acinus results in cells displaying morphological changes associated with apoptosis. 68
Figure 3.11. Overexpression of Acinus causes apoptosis. 70
Figure 3.12. Overexpression of Acinus decreases cell viability. 71
Figure 4.1. CLIP (Cross-Linking and Immuno-Precipitation) assay. 78
Figure 4.2. Analysis of Acinus-RNA complexes by CLIP. 81
Figure 4.3. RNA targets of Acinus identified by CLIP which are located in or nearby sites of alternative splicing events. 84
Figure 4.4. Analysis of Acinus knockdown and overexpression. 92
Figure 4.5. Quality analysis of RNA to be used in RT-PCR assays. 93
Figure 4.6. Alternative splicing regulation of Acinus target transcripts. 96
Figure 5.1. Analysis of hnRNP A1-RNA complexes by CLIP. 104
Figure 5.2. Analysis of transcripts bound by hnRNP A1 with respect to their experimental identification. 105
Figure 5.3. RNA targets of hnRNP A1 identified by CLIP which are located in or nearby sites of alternative splicing events. 109
Figure 5.4. HnRNP A1 CLIP tags which contain sequences similar to the SELEX winner sequence UAGGGA/U. 115
Figure 5.5. Identification of the hnRNP A1 binding site. 116
Figure 5.6. HnRNP A1 is recruited to stress granules (SGs) in response to osmotic shock (OSM). 117

Figure 5.7. Analysis of hnRNP A1 knockdown and overexpression in the presence or absence of osmotic shock. 118
Figure 5.8. Quality analysis of RNA to be used in RT-PCR assays. 119
Figure 5.9. Alternative splicing regulation of hnRNP A1 target transcripts. 124
Figure 5.10. RNA binding proteins exhibit distinct binding site location preferences. 130

List of Tables

Table 1.1. “Classical” SR proteins. 13
Table 1.2. Additional SR proteins. 14
Table 1.3. RNA-binding SR related factors. 14
Table 1.4. Other RS domain containing proteins. 15
Table 1.5. RNA sequences identified as SR protein binding sites. 22
Table 2.1. Preparation of Materials. 41
Table 2.2. Primer sequences used in RT-PCR alternative splicing assays. 52
Table 4.1. Optimisation of Acinus CLIP. 80
Table 4.2. Alternative cassette exons bound by Acinus. 83
Table 4.3. Acinus target exons containing alternative 3' splice sites. 83
Table 4.4. Acinus target exons containing alternative 5' splice sites. 83
Table 4.5. Acinus target exons containing a retained intron. 83
Table 4.6. Acinus target exons located adjacent to alternative cassette exons. 83
Table 4.7. Acinus target introns located adjacent to alternative cassette exons. 83
Table 4.8. Effect on protein product of alternative splicing of Acinus bound transcripts. 86
Table 4.9. Other Acinus target transcripts. 87
Table 4.10. Analysis of protein-coding Acinus CLIP tag locations in relation to exon-exon junctions. 89
Table 5.1. Alternative cassette exons bound by hnRNP A1. 107
Table 5.2. HnRNP A1 target exons containing alternative 3' splice sites. 107
Table 5.3. HnRNP A1 target exons containing alternative 5' splice sites. 107
Table 5.4. HnRNP A1 target exons containing a retained intron. 107
Table 5.5. HnRNP A1 target exons located adjacent to alternative cassette exons. 108
Table 5.6. HnRNP A1 target introns located adjacent to alternative cassette exons. 108
Table 5.7. Effect on protein product of alternative splicing of hnRNP A1 bound transcripts. 112
Table 5.8. Other hnRNP A1 target transcripts. 113

Chapter 1: Introduction

1.1. Gene expression and RNA processing

Gene expression involves the transmission of information encoded within the sequence of nucleotides of DNA into a functional protein product. Eukaryotic gene expression begins with a DNA template, which is transcribed in the nucleus by RNA Polymerase II (RNAP II) to an RNA transcript. Three main processing events must occur before the RNA can be exported to the cytoplasm as a translatable mRNA. These are 5' capping, 3' end formation and RNA splicing. Although each of these events are biochemically distinct from one another, they can influence one another's specificity and efficiency, and are all tightly linked to transcription. The C-terminal domain (CTD) of RNAP II is thought to mediate the coupling of these processes by acting as a protein scaffold (Maniatis and Reed, 2002).

The CTD is unique to RNAP II, it is not found in either of the other RNA polymerases, which are responsible for the transcription of pre-rRNAs and pre-tRNAs. The CTD is comprised of multiple repeats of the consensus sequence YSPTSPS, which is highly conserved throughout evolution and is subject to reversible phosphorylation during the transcription cycle. Binding of RNA processing factors to the CTD is dependent upon its phosphorylation status. The enzymes responsible for cap synthesis, RNA guanylyltransferase and RNA (guanine-7)-methyltransferase, bind directly to the phosphorylated, but not to the non-phosphorylated form of the CTD *in vitro* (McCracken *et al.*, 1997a). The cleavage and polyadenylation factors, CPSF and CstF, also bind RNAP II via the CTD and co-purify with RNAP II in a high molecular mass complex (McCracken *et al.*, 1997b.). Likewise, splicing factors have been shown to bind to the phosphorylated CTD of RNAP II (Corden and Patturajan, 1997; Morris and Greenleaf, 2000). Splicing, 3' end processing and termination of transcription downstream of the poly(A) site are all inhibited upon truncation of the CTD (McCracken *et al.*, 1997b). The crystal structure of RNAP II shows that the nascent RNA exits RNAP II in the vicinity of the CTD (Cramer *et al.*, 2001), where bound processing factors can act upon the RNA, thereby coupling transcription, capping, splicing and polyadenylation.

1.1. Pre-mRNA splicing

Pre-mRNA splicing was discovered in the late 1970s when it was demonstrated that eukaryotic genes contained intervening sequences, or introns, that were not present in the mature mRNA (Berget *et al.*, 1977; Chow *et al.*, 1977).

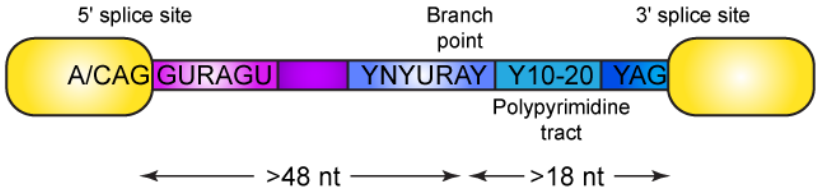


Figure 1.1. Splice site elements (adapted from Smith and Vacarcel, 2000). Consensus splice site elements for a typical metazoan intron (R, purine; Y, pyrimidine; N, any nucleotide). The minimal distances between the 5' splice site and branch point, and between the 3' splice site and branch point are indicated.

Pre-mRNAs contain essential consensus sequence elements in cis, which direct RNA splicing (Figure 1.1). The junction between the exon and adjacent downstream intron, or 5' splice site (ss), is marked by the consensus AG|GURAGU (R, purine; Y, pyrimidine; | indicates the exon-intron junction). While the 3'ss at the end of the intron is defined by YAG|R. The branch point (BP), YNYURAY, containing a highly conserved adenosine, is found upstream of the 3'ss preceding a polypyrimidine tract. These elements are necessary for the chemical reaction of splicing to proceed. This occurs via two transesterification reactions (Figure 1.2). Nucleophilic attack of the 5'ss by the 2'OH of the BP adenosine and the first transesterification reaction results in a free 5' exon and a lariat containing the intron sequences plus the 3' exon. The 3'OH of the freed 5' exon then attacks the 3'ss and the ensuing transesterification reaction results in fusion of the 5' and 3' exons and release of the lariat-shaped intron (Moore and Sharp, 1993).

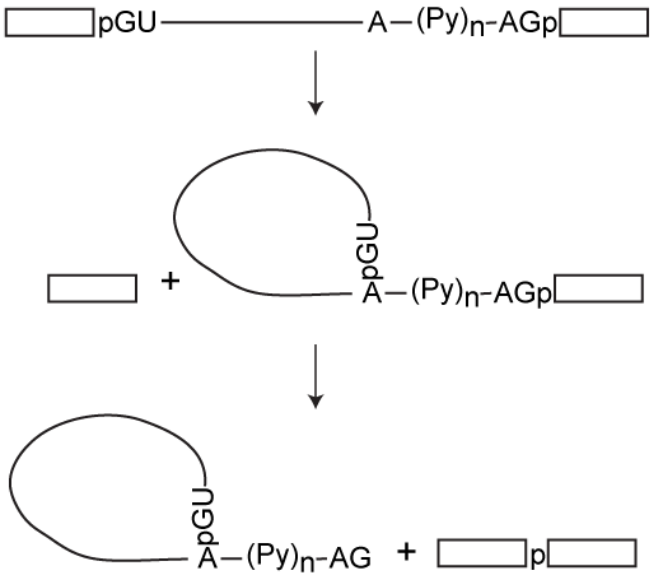


Figure 1.2. Splicing takes place in two transesterification reactions (adapted from Black, 2003). The first reaction results in a free 5' exon and the intron and 3' exon contained in a lariat structure. After the second transesterification reaction the two exons are ligated together and the intron is released as a lariat.

1.2. The spliceosome

Introns are removed by a macromolecular complex, termed the spliceosome, which consists of five small nuclear ribonucleoprotein particles (snRNPs) U1, U2, U4, U5 and U6, and a large number of protein components (Kramer, 1996; Will and Luhrmann, 2001). The majority of pre-mRNAs are spliced by the abundant U2-dependent (major) spliceosome, which will be discussed below. However, it is important to note that a rare U12-type class of pre-mRNA introns are spliced by the less abundant U12-dependent (minor) spliceosome (Sharp and Burge, 1997).

Spliceosomal assembly upon the pre-mRNA is an ordered process and distinct intermediate complexes can be detected (Figure 1.3). Assembly is initiated by recognition of the 5' and 3' splice sites by the U1 snRNP and the heterodimeric U2 snRNP Auxiliary Factor (U2AF), respectively, forming the E complex. Recruitment of the U2 snRNP to the BP, in an ATP-dependent manner, results in the formation of the A complex. In addition, this process bulges out the BP adenosine, enhancing the first nucleophilic attack. The U2-branch point interaction is also facilitated by protein-pre-mRNA interactions involving SF3b and SF3a subunits, as well as the

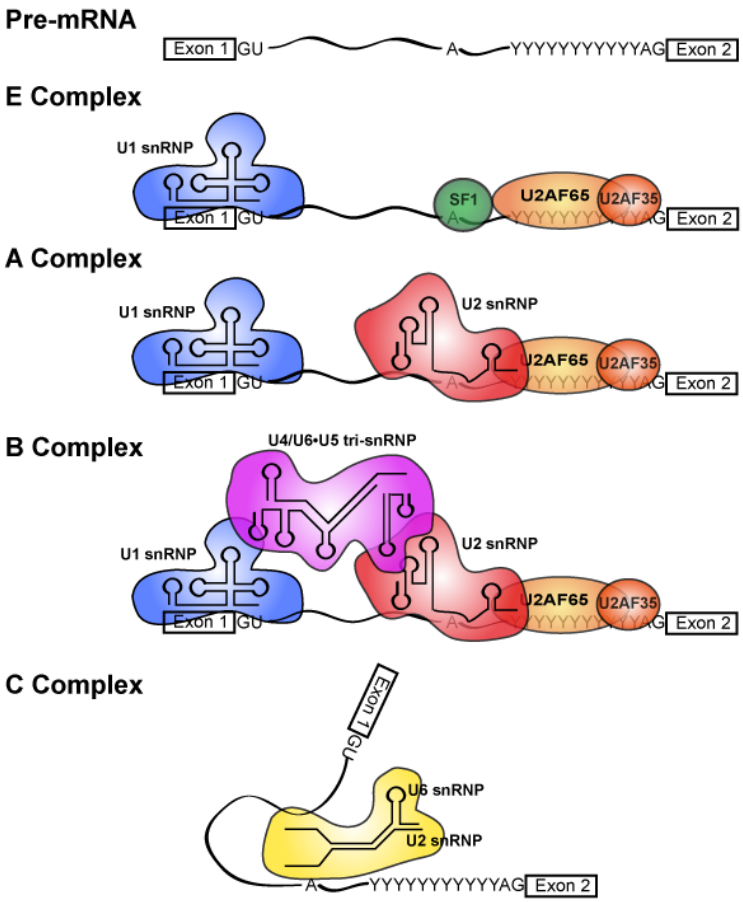


Figure 1.3. Spliceosome assembly (adapted from Hertel and Graveley, 2005). Spliceosome assembly upon the pre-mRNA is an ordered process. The E complex contains U1 snRNP bound to the 5' splice site, SF1 bound to the branch point and U2AF65 and U2AF35 bound to the polypyrimidine tract and the 3' splice site AG, respectively. SF1 is replaced by U2 snRNP in the A complex, and the B complex is formed after the association of the U4/U6•U5 tri-snRNP. After massive structural rearrangements the catalytically active C complex is formed.

non-snRNP splicing factors U2AF and SF1/BBP (Gozani *et al.*, 1996; Gozani *et al.*, 1998). Subsequent recruitment of the U4/U6·U5 tri-snRNP forms the B complex, which is followed by a series of structural rearrangements leading to the formation of the catalytically active spliceosomal C complex (reviewed by (Matlin and Moore, 2007)). This remodelling involves displacement of the U1 snRNP from the 5'ss by U6 snRNA base-pairing to the 5'ss sequence, and disruption of the U4/U6 interaction by U6 snRNA base-pairing to U2 snRNA. These complex RNA-RNA interactions between U2, U6 and the pre-mRNA form a catalytic core, which position the 5'ss and BP in close proximity (Madhani and Guthrie, 1992; Sun and Manley, 1995).

The consensus model of spliceosome assembly (E→A→B→C complex) is described above. However, some components of the spliceosome have been shown to

associate to the pre-mRNA at an earlier step. For example, 17S U2 snRNPs have been shown to be functionally associated with the pre-mRNA at the time of E complex formation (Das *et al.*, 2000). It appears that the U2 snRNP first binds loosely to the pre-mRNA in the E-complex, via the integral U2 snRNP-associated protein SF3b, before the ATP-dependent stable binding of the U2 snRNP to the BP occurs in the A complex.

The U4/U6•U5 tri-snRNP has also been reported to enter the spliceosome at an earlier stage than in complex B of the consensus model. The U4/U6•U5 tri-snRNP recognises the 5'ss, together with U1, at the early stages of spliceosomal assembly (Maroney *et al.*, 2000). This recognition appears to be ATP-dependent and can occur in the absence of a stable U2 snRNP-BP interaction. Both U1 snRNP and the U4/U6•U5 tri-snRNP define the 5'ss before interactions with the 3'ss and U2 snRNP occur. The conserved loop I of U5 snRNA can also interact with both 5' and 3' exon sequences (Newman and Norman, 1992; Sontheimer and Steitz, 1992; Wyatt *et al.*, 1992; Cortes *et al.*, 1993; Newman *et al.*, 1995). The U5 loop can therefore 'tether' the 5' exon splicing intermediate produced by the first transesterification, and then assist in the alignment of the 5' and 3' exons for the second catalytic step (Newman and Norman, 1992; Sontheimer and Steitz, 1992). Recognition of splicing elements by multiple factors most likely aids splicing fidelity. In addition, RNAP II helps to commit the nascent RNA to splicing and ensures the proper pairing of splice sites by facilitating the binding of U1 and/or U2 snRNPs to the 5'ss and/or BP (Hirose *et al.*, 1999)..

The spliceosome is a dynamic structure and more than 300 proteins have been identified in active splicing complexes (Rappsilber *et al.*, 2002; Zhou *et al.*, 2002; Bessonov *et al.*, 2008; reviewed by (Jurica and Moore, 2003)). Despite this there is growing evidence that the catalytic component of the spliceosome is RNA based. Purified U2 and U6 RNA fragments can form intermolecular base pairing, similar to that seen in the spliceosome, in the absence of any protein factors (Valadkhan and Manley, 2000). Furthermore, this protein-free complex of U2 and U6 snRNAs can bind and position a small RNA containing the BP sequence. The reaction that ensues is related to the first step of splicing (Valadkhan and Manley, 2001). In addition, it

has been reported that metal ion co-ordination by yeast U6 snRNA is necessary for splicing (Yean *et al.*, 2000; Valadkhan and Manley, 2002). However, the protein components of the spliceosome are essential for the assembly and stabilisation of the spliceosome *in vivo*.

1.3. The Exon Junction Complex

Pre-mRNA splicing can 'mark' mRNAs with information that affects their subsequent metabolism, illustrating another point of control in the complex web of interconnected processes involved in gene expression. This is achieved by altering the complement of proteins associated with the mRNA. Proteins that specifically associate with spliced RNAs were identified by crosslinking and immunoprecipitation approaches (Kataoka *et al.*, 2000; Le Hir *et al.*, 2000a). Subsequently it was demonstrated these proteins are deposited at a specific position about 20-24 nts upstream of exon-exon junctions (Le Hir *et al.*, 2000b; Le Hir *et al.*, 2003). This set of proteins is collectively known as the exon junction complex (EJC).

The EJC is a dynamic structure consisting of a few core proteins and several more peripheral components (Figure 1.4). The EJC core proteins provide a binding platform for the transiently associated factors, which include splicing-related proteins, mRNA export factors and proteins involved in nonsense mediated decay (NMD). This allows the EJC to act in, and facilitate, several nuclear and cytoplasmic mRNA processing events (Tange *et al.*, 2004) including subcellular localisation and regulated translation of specific mRNAs (reviewed by (Giorgi and Moore, 2007)), mediation of the enhancement effect of splicing on translation (Wiegand *et al.*, 2003; Nott *et al.*, 2004) and NMD (Kashima *et al.*, 2006; Chang *et al.*, 2007). In mammalian cells NMD requires the presence of a spliceable intron downstream of a premature termination codon (PTC) (Cheng *et al.*, 1994; Carter *et al.*, 1996; Thermann *et al.*, 1998). Therefore splicing leaves a 'mark' that signals the positions of exon-exon boundaries to the translation and decay machineries (Li and Wilkinson, 1998; Nagy and Maquat, 1998; Wagner and Lykke-Andersen, 2002).

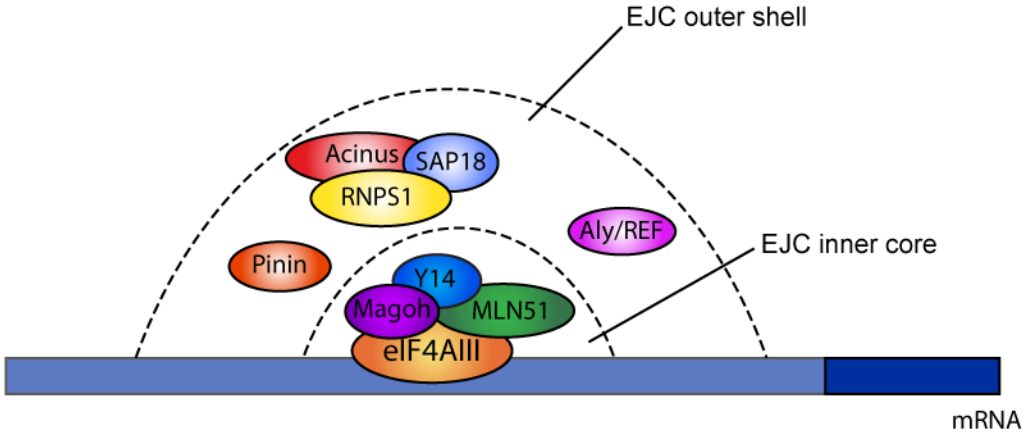


Figure 1.4. Protein components of the exon junction complex (EJC) (adapted from Tange *et al.*, 2005). The minimal EJC core consists of a tetrameric complex containing eIF4AIII, MLN51, Magoh, and Y14, with eIF4AIII providing direct contact to the mRNA. These core proteins are nucleocytoplasmic shuttling proteins and most likely follow the mRNA to the cytoplasm. Proteins in the outer shell were identified in the purified EJC. RNPS1, Acinus, and SAP18 can stably associate and may bind the EJC core as a trimeric complex. However, RNPS1 may also bind alone via interactions with Pinin. SAP18, RNPS1, and Aly/REF are shuttling proteins whereas Acinus and Pinin are nuclear restricted.Transiently interacting factors form a third layer of the EJC, these are proteins not identified in the in vitro-derived EJC, but which are likely to interact dynamically with either the EJC core or outer sphere proteins.

mRNA binding is mediated by the DEAD-box protein eIF4AIII (Shibuya *et al.*, 2004; Le Hir and Andersen, 2008). Inhibition of its ATPase activity allows the EJC to form a stable association with the mRNA (Ballut *et al.*, 2005). This association is independent of the mRNA sequence (Le Hir *et al.*, 2000b). Therefore an EJC is expected to be loaded on every spliced junction of mature mRNAs. The complex is assembled coincidentally with the second step of splicing leading to exon ligation (Reichert *et al.*, 2002; Tange *et al.*, 2005; Merz *et al.*, 2007), and remains during nuclear export and cytoplasmic localisation (Kim *et al.*, 2001; Le Hir *et al.*, 2001) until it is removed in the first translation round (Dostie and Dreyfuss, 2002; Lejeune *et al.*, 2002).

1.4. Alternative splicing

Following the initial sequencing and analysis of the human genome in 2001 (Lander *et al.*, 2001; Venter *et al.*, 2001), the realisation there are fewer human genes than anticipated caused an increased interest in alternative splicing. This process enables the generation of distinct mRNAs from a common precursor. This violates the “one

gene, one polypeptide" rule, and was believed to affect only a minority of genes. In 1993 Phil Sharp speculated in his Nobel Prize lecture that about 5% of human genes would be alternatively spliced (Sharp, 1994). However, bioinformatic analyses of expressed sequence tags (ESTs) conservatively estimated that about 35%-60% of human genes were alternatively spliced (Modrek *et al.*, 2001). A further study used exon-junction microarrays to monitor splicing at every exon-exon junction in more than 10,000 multi-exon human genes in 52 tissues and cell lines (Johnson *et al.*, 2003). When the number of alternative splicing events detected in this study was added to those from the EST analysis, the estimate of human genes that were alternatively spliced increased to 73%. These studies highlight the large number of human genes which are subject to regulation by alternative splicing.

Alternative splicing can generate numerous transcripts from a single gene by joining together different combinations of splice sites (Figure 1.5). The Drosophila Dscam gene exemplifies the extreme structural diversity achievable by alternative splicing. Dscam, encoding the Down syndrome cell-adhesion molecule, is required for the development of neural circuits. There are four blocks of alternatively used exons in the Dscam gene: the exon-4 block contains 12 alternatives; the exon-6 block contains 48 alternatives; the exon-9 block contains 33 alternatives; and the exon-17 block contains 2 alternatives. This gives a potential of 38,016 different isoforms that can be generated by alternative splicing from one gene (Schmucker *et al.*, 2000). From this example it is clear that alternative splicing must be highly regulated in order to ensure the correct isoforms are produced in specific cell-types, developmental stages or signalling pathways (reviewed by (Matlin *et al.*, 2005)).

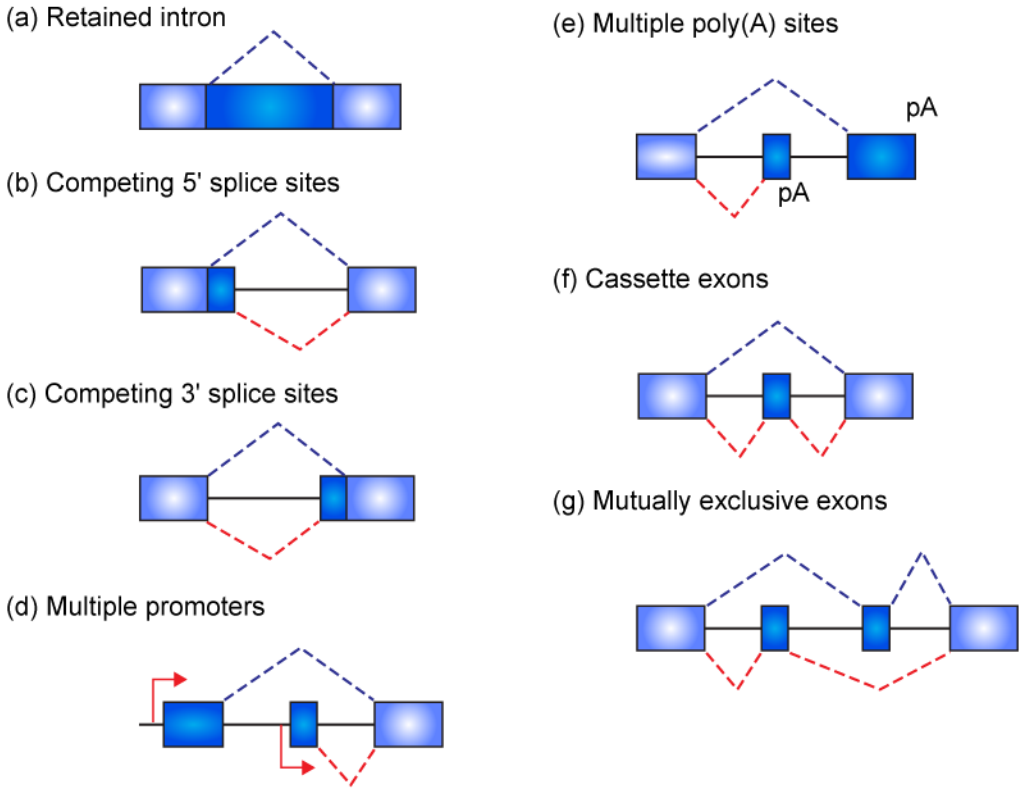


Figure 1.5. Elementary alternative splicing events (adapted from Roberts and Smith, 2002). (a) a sequence can be excised as an intron or remain in the mRNA. Individual exons can be lengthened or shortened by the use of alternative 5' (b) or 3' (c) splice sites. At the 5' and 3' ends of genes, alternative splicing can occur in association with selection of (d) alternative promoters, shown as arrows, or (e) polyadenylation sites, indicated by 'pA'. (f) Internal cassette exons can be included or skipped independently of other exons, while (g) mutually exclusive splicing involves an array of two or more alternative exons, only one of which can be included in the mature mRNA.

Promoter choice, under transcriptional control, can also influence the splicing pattern adopted. However, the regulation of alternative splicing is predominantly due to the composition of the spliceosome and the roles of other non-spliceosomal trans-acting factors. Splice site choice is thought to be regulated by altering the binding of initial factors to the pre-mRNA and the formation of the early spliceosomal complexes. By the time the E complex is formed, it appears that the splice sites are paired in a functional sense and the defined intron is committed to being spliced. However, there is evidence that regulation of the 3'ss can occur after the first transesterification reaction (Lallena *et al.*, 2002).

1.5. The SR protein family

The SR proteins were first discovered as splicing factors in the early 1990s (Fu, 1995; Graveley, 2000; Bourgeois *et al.*, 2004). A protein domain rich in arginine and serine dipeptides, termed the RS domain, was originally observed in three *Drosophila* splicing regulators, *suppressor-of-white-apricot* (SWAP) (Chou *et al.*, 1987), *transformer* (Tra) (Boggs *et al.*, 1987) and *transformer-2* (Tra-2) (Amrein *et al.*, 1988). Subsequent identification of SF2/ASF (Splicing factor 2/ Alternative Splicing Factor) (Ge *et al.*, 1991; Krainer *et al.*, 1991) and SC35 (spliceosomal component 35) (Fu and Maniatis, 1990) revealed that these proteins contained an RS domain, which is also present in the U1 snRNP-associated protein, U1-70K (Theissen *et al.*, 1986; Spritz *et al.*, 1987).

SF2/ASF was the first SR protein to be identified as an activity required to complement an otherwise splicing-deficient HeLa S100 extract (Krainer *et al.*, 1990a) and was also purified from 293 cells as a factor which could alter 5'ss selection of an SV40 early pre-mRNA (Ge and Manley, 1990). The term 'SR protein' was coined following identification of additional RS domain-containing proteins that were recognised by a monoclonal antibody, mAb 104, which binds to active sites of RNAP II transcription (Roth *et al.*, 1990). These novel proteins, which were active in splicing complementation, included the SR proteins SRp20, SRp40, SRp55 and SRp75, named after their apparent molecular mass on an SDS-PAGE gel and are conserved across vertebrates and invertebrates (Zahler *et al.*, 1992). They have a modular structure containing one or two copies of an RRM (RNA recognition motif) at the N-terminus that provides RNA-binding specificity and a C-terminal RS domain that acts to promote protein-protein interactions that facilitate recruitment of the spliceosome (Wu and Maniatis, 1993; Kohtz *et al.*, 1994). The RS domain can also contact the pre-mRNA directly via the BP and the 5'ss suggesting an alternative way to promote spliceosome assembly (Shen and Green, 2004; Shen *et al.*, 2004). Furthermore, the RS domain acts as a NLS (nuclear localisation signal), affecting the subcellular localisation of SR proteins by mediating the interaction with the SR protein nuclear import receptor, transportin-SR (Caceres *et al.*, 1997; Kataoka *et al.*, 1999; Lai *et al.*, 2000).

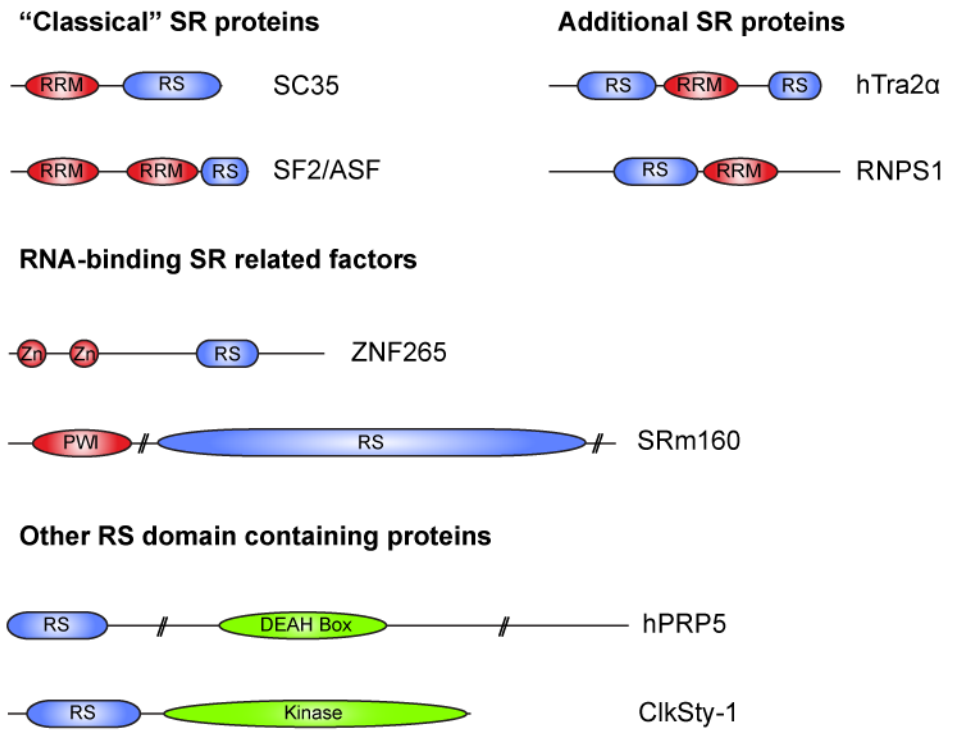


Figure 1.6 Schematic diagram of SR and SR-related proteins. The domain structures are depicted. RRM: RNA recognition motif; RS: arginine/serine-rich domain; Zn: zinc finger motif; PWI: an alternative RNA binding motif; DEAH Box: motif characteristic of RNA helicases. With the exception of SRm160 and hPRP5, all proteins are drawn to scale.

The prototypical SR protein, SF2/ASF, functions in constitutive splicing and also modulates alternative splicing (Ge and Manley, 1990; Krainer *et al.*, 1990b). Further studies demonstrated that other SR protein family members could also affect alternative splicing *in vitro* (Fu *et al.*, 1992; Zahler *et al.*, 1993a). Thus, the criteria used to define classical SR protein family members are i) structural similarity ii) dual function in constitutive and alternative splicing iii) the presence of a phosphoepitope recognised by mAb104 iv) their purification using magnesium chloride (Table 1.1).

A genome-wide survey in metazoans identified a large number of RS domain-containing proteins with a role not only in splicing but also in other cellular processes such as chromatin remodelling, transcription and cell cycle progression (Boucher *et al.*, 2001). These related proteins contain an RS domain but may lack a defined RRM, however, a subset can bind RNA through other domains such as the PWI motif found in the splicing activator SRm160 (Blencowe *et al.*, 1998; Szymczyna *et al.*, 2003). These factors are collectively known as SR-like or SR-related proteins and include both subunits of the U2AF heterodimer, U1-70K and the

splicing coactivators SRm160/300 among others (Blencowe *et al.*, 1999). It was recently proposed that SR proteins should be redefined based on their common structural features and their role in pre-mRNA splicing (Lin and Fu, 2007). Based on this, a 'bona fide' SR protein has to contain at least one RRM and an RS domain (irrespective of their position within the protein) and to function in constitutive or alternative splicing, as assayed by either complementation of splicing-deficient S100 HeLa cytoplasmic extracts or in an alternative splicing assay, respectively. The human homologues of the *Drosophila* splicing regulators Tra2α (Dauwalder *et al.*, 1996) and Tra2β (Beil *et al.*, 1997) contain an RRM flanked by two RS domains, which is not the domain structure found in classical SR proteins (Figure 1.6). Since both proteins function as sequence-specific splicing activators (Tacke *et al.*, 1998), they could be classified as SR proteins (Table 1.2). Other proteins may contain an RS domain but also have other domains required for their enzymatic activities, as is the case with the RNA helicases HRH1 and hPRP16, which both contain a DEAH box domain in addition to an RS domain (Ohno and Shimura, 1996; Zhou and Reed, 1998) (Table 1.4).

Table 1.1. "Classical" SR proteins

Protein name	Gene name	Key domains	Splicing role	UniProt
SF2/ASF	SFRS1	Two RRMs, RS domain	Constitutive and alternative splicing activator	Q07955
SC35	SFRS2	RRM, RS domain	Constitutive and alternative splicing activator	Q01130
SRp20	SFRS3	RRM, RS domain	Constitutive and alternative splicing activator	P84103
SRp75	SFRS4	Two RRMs, RS domain	Constitutive and alternative splicing activator	Q08170
SRp40	SFRS5	Two RRMs, RS domain	Constitutive and alternative splicing activator	Q13243
SRp55	SFRS6	Two RRMs, RS domain	Constitutive and alternative splicing activator	Q13247
9G8	SFRS7	RRM, RS domain, CCHC-type Zinc finger	Constitutive and alternative splicing activator	Q16629

Table 1.2. Additional SR proteins

Protein name	Gene name	Key domains	Splicing role	UniProt
SRp54	SFRS11	RRM, RS domain	Alternative splicing repressor	Q05519
SRp30c	SFRS9	Two RRMs, RS domain	Constitutive and alternative splicing regulator	Q13242
SRp38, TASR	FUSIP1	RRM, RS domain	General splicing repressor	O75494
hTra2α	TRA2A	RRM, two RS domains	Splicing activator	Q13595
hTra2β	SFRS10	RRM, two RS domains	Splicing activator	P62995
RNPS1	RNPS1	RRM, RS domain	Constitutive and alternative splicing regulator	Q15287
SRrp35	SRRP35	RRM, RS domain	Negative regulator of alternative splicing	Q8WXF0
SRrp86, SRrp508	SFRS12	RRM, RS domain	Positive and negative regulator of alternative splicing	Q8WXA9
U2AF35	U2AF1	RRM, RS domain, two C3H1 type zinc fingers	Essential constitutive splicing factor	Q01081
U2AF65	U2AF2	Three RRMs, RS domain	Essential constitutive splicing factor	P26368
U1-70k	SNRP70	RRM, RS domain	Essential constitutive splicing factor	P08621
XE7	SFRS17A	RRM, RS domain	Alternative splicing regulator	Q02040
SRp46	SFRS2B	RRM, RS domain	Constitutive and alternative splicing regulator	Q9BRL6

Table 1.3. RNA-binding SR related factors

Protein name	Gene name	Key domains	Splicing role	UniProt
Urp	ZRSR2	RRM, RS domain	Splicing factor	Q15696
HCC1/CAPER	RBM39	RRM, RS domain	Alternative splicing regulator	Q14498
hSWAP	SFRS16	RS domain	Alternative splicing regulator	Q8N2M8
Pinin	PNN	RS domain	Alternative splicing regulator	Q9H307
SRrp129	SFRS2IP	RS domain	Splicing factor	Q99590
U4/U6.U5 tri-snRNP-associated 27 kDa protein	RY-1	RS domain	Unknown	Q8WVK2
LUC7B1	LUC7L	RS domain, C2H2 type zinc finger	Unknown	Q9NQ29
Acinus	ACIN1	RRM, RS domain, SAP domain	Unknown	Q9UKV3
SR-A1	SFRS19/ SCAF1	RS domain	Unknown	Q9H7N4
ZNF265	ZRANB2	RS domain, two RANBP2 type zinc fingers	Alternative splicing regulator	O95218
SRm160	SRRM1	RS domain, PWI domain	Constitutive and alternative splicing co-activator	Q8IYB3
SRm300	SRRM2	RS domain	Constitutive and alternative splicing co-activator	Q9UQ35
RBM5	RBM5	Two RRMs, RS domain, RANBP2 type zinc finger, C2H2 type zinc finger	Unknown	P52756
U2 associated protein SR140	SR140	RRM, RS domain	Unknown	O15042
RBM23	RBM23	Two RRMs, RS domain	Unknown	Q86U06
SFRS15	SFRS15	RRM, RS domain	Unknown	O95104

Table 1.4. Other RS domain containing proteins

Protein name	Gene name	Key domains	Splicing role	UniProt
SRrp53	RSRC1	RS domain, coiled-coil domain	Unknown	Q96IZ7
hPRP5	DDX46	RS domain, DEAH box	Spliceosomal rearrangement	Q7L014
hPRP16	DHX38	RS domain, DEAH box	Splicing factor	Q92620
Prp22/HRH1	DHX8	RS domain, DEAH box	Spliceosomal rearrangement	Q14562
U5-100k/ hPRP28	DDX23	RS domain, DEAD box	Spliceosomal rearrangement	Q9BUQ8
ClkSty-1	CLK1	RS domain, kinase domain	SR protein kinase	P49759
ClkSty-2	CLK2	RS domain, kinase domain	SR protein kinase	P49760
ClkSty-3	CLK3	RS domain, kinase domain	SR protein kinase	P49761
Prp4k	PRPF4B	RS domain, kinase domain	SR protein kinase	Q13523
CrkRS	CRKRS	RS domain, kinase domain	SR protein kinase	Q9NYV4
CDC2L5	CDC2L5	RS domain, kinase domain	Alternative splicing regulator	Q14004
Cyclin-L1	CCNL1	RS domain, two cyclin-like domains	Alternative splicing regulator	Q9UK58
Cyclin-L2	CCNL2	RS domain, two cyclin-like domains	Alternative splicing regulator	Q96S94
SR-cyp	PPIG	RS domain, PPIase cyclophilin-type domain	Regulates localisation of SR proteins	Q13427
CIR	CIR	RS domain	Alternative splicing regulator	Q86X95
SRrp130	SFRS18	Two RS domains	Unknown	Q8TF01

1.6. Co-transcriptional splicing

Splicing factors are concentrated in nuclear speckles and are recruited from these sites to nascent sites of RNAP II transcription (Misteli *et al.*, 1997). It is well documented that RNA splicing occurs co-transcriptionally (Bauren and Wieslander, 1994; Beyer and Osheim, 1988). Interactions between SR-related proteins and the CTD of RNAP II have been reported (Yuryev *et al.*, 1996), and members of the SR protein family were identified among the hundreds of proteins present in the RNAP II complex (Das *et al.*, 2007). It was recently reported that SC35 promotes RNAP II elongation in a subset of genes confirming the existence of coupling between transcription and splicing, and perhaps surprisingly, showing that this coupling can be bidirectional (Lin *et al.*, 2008). In this study it was demonstrated that SC35 interacts not only with the CTD but also with CDK9, the kinase component of the transcriptional elongation factor P-TEFb, resulting in phosphorylation of Ser2 in the CTD and leading to transcriptional elongation. This activity of SR proteins in

transcriptional elongation may be functionally related to their reported effect in the maintenance of genome stability. It has been shown that depletion of SF2/ASF, SC35 and the SR-related protein RNSP1 results in the formation of R-loops (RNA:DNA hybrid structures) leading to a hypermutation phenotype (Li and Manley, 2005; Li *et al.*, 2007; Xiao *et al.*, 2007).

The co-transcriptional nature of pre-mRNA splicing underlies a role for the transcriptional machinery in alternative splicing regulation (Hicks *et al.*, 2006). A kinetic coupling model proposed that changes in the rate of transcriptional elongation affect the timing in which splice sites are presented to the splicing machinery, leading to differential splice site selection (de la Mata *et al.*, 2003). Furthermore, differential recruitment of splicing factors to the CTD of RNAP II may also influence this process (de la Mata and Kornblihtt, 2006; reviewed by (Kornblihtt, 2007)).

1.7. The heterogeneous nuclear ribonucleoparticle (hnRNP) family

The heterogeneous nuclear ribonucleoparticle (hnRNP) family comprises a structurally diverse group of RNA-binding proteins. They were first described as a major group of chromatin-associated RNA-binding proteins (Dreyfuss *et al.*, 1993), which were thought to bind nascent RNA in a non-specific manner and package hnRNAs into hnRNP particles (Sun *et al.*, 1998). However, structural and functional analyses showed individual hnRNP proteins contained RNA-binding motifs with different RNA sequence binding preferences (Dreyfuss *et al.*, 1993).

HnRNPs, like SR proteins, are widely expressed, associate with nascent transcripts and have roles in many aspects of RNA biogenesis, including pre-mRNA splicing (Martinez-Conteras *et al.*, 2007), nucleocytoplasmic transport of mRNA (Izaurralde *et al.*, 1997), mRNA localisation (Hoek *et al.*, 1998; Cote *et al.*, 1999), translation (Ostareck *et al.*, 1997; Collier *et al.*, 1998; Habelhah *et al.*, 2001) and mRNA stability (Kiledjian *et al.*, 1997; Xu *et al.*, 2001). SR proteins and hnRNP proteins effectively belong to the same protein family, with both sets of proteins exhibiting similar properties and functions.

HnRNP A1 has been extensively studied and is a very abundant nuclear protein. It is a regulator of alternative splicing and acts by antagonising the SR family of proteins as discussed below (Mayeda and Krainer, 1992; Caceres *et al.*, 1994; Yang *et al.*, 1994). HnRNP A1 contains two RRMs at the N-terminus through which it binds RNA. An RGG (Arg-Gly-Gly) box at the C-terminus also contributes to RNA binding (Mayeda *et al.*, 1994) (Figure 1.7). Although hnRNP A1 is primarily nuclear, it shuttles continuously between the nucleus and the cytoplasm (Pinol-Roma and Dreyfuss, 1992), which allows hnRNP A1 to function in both nuclear and cytoplasmic processes. This nucleocytoplasmic shuttling is conferred by a 38 amino acid domain, M9, found at the C-terminus (Michael *et al.*, 1995; Siomi and Dreyfuss, 1995; Izaurralde *et al.*, 1997) (Figure 1.7).

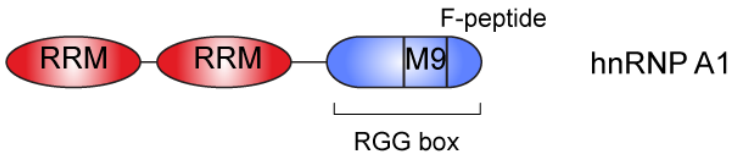


Figure 1.7 Features of hnRNP A1. The two RRM domains at the N-terminus are responsible for RNA binding, as is the RGG box at the C-terminus. Within the RGG box is the M9 motif necessary for hnRNP A1 nuclear export and import, and the F-peptide which is phosphorylated in response to stress.

1.8. Roles of classical SR proteins in constitutive and alternative splicing

Splice site consensus sequences are generally not sufficient to direct assembly of a functional spliceosome, and auxiliary elements known as ESEs and ISEs (exonic and intronic splicing enhancers, respectively) and ESSs and ISSs (exonic and intronic splicing silencers, respectively) are involved in both constitutive and, to perhaps a greater extent, alternative splicing. Binding of SR proteins to ESEs acts as a barrier that prevents exon skipping, thus ensuring the correct 5' to 3' linear order of exons in spliced mRNA (Ibrahim *et al.*, 2005). Two main models have been proposed to explain the mechanism by which SR proteins regulate exon inclusion. The "recruitment model" focuses on the ability of ESE-bound SR proteins to recruit and stabilise interactions between the U1 snRNP at the 5'ss and U2AF65 at the 3'ss

(Graveley *et al.*, 2001), in a process known as exon definition (Robberson *et al.*, 1990) (Figure 1.8A). In the "inhibitor model" ESE-bound SR proteins may act by antagonising the negative activity of hnRNP proteins recognising ESSs (Zhu *et al.*, 2001) (Figure 1.8B). The SR proteins may also form a network of protein-protein interactions across introns to juxtapose the 5' and 3'ss early in spliceosomal assembly, as shown by the reported interactions of SF2/ASF and SC35 with U1-70K at the 5'ss and with U2AF35 at the 3'ss in a RS domain-dependent manner (Wu and Maniatis, 1993) (Figure 1.8C). Additionally, the enhancer-bound RS domain of the SR protein SF2/ASF has been shown to directly interact with the RNA at the BP to promote pre-spliceosomal assembly (Shen and Green, 2004; Shen *et al.*, 2004). SR proteins may also facilitate the recruitment of the U4/U6•U5 tri-snRNP to the pre-spliceosome (Roscigno and Garcia-Blanco, 1995) via RS domain-mediated interactions with the SR-related proteins SRrp65 and SRrp110 (Makarova *et al.*, 2001). The function of SF2/ASF in pre-mRNA splicing depends on the context of the pre-mRNA sequence to which it binds, as shown by the fact that SF2/ASF inhibits adenovirus IIIa pre-mRNA splicing when bound to an intronic repressor element (Kanopka *et al.*, 1996). The second RRM of SF2/ASF, and in particular a phylogenetically conserved heptapeptide, SWQDLKD, which is located in the first α-helix of this domain (Birney *et al.*, 1993), is essential for the splice site selection activity of SF2/ASF (Dauksaite and Akusjarvi, 2002; Dauksaite and Akusjarvi, 2004).

Use of fluorescence recovery after photobleaching (FRAP) approaches revealed a high mobility for SF2/ASF within the nucleus with kinetics compatible with a diffusion mechanism (Kruhlak *et al.*, 2000; Phair and Misteli, 2000). Advances in imaging have allowed analysis of splicing factors both in speckles and at other sites in the nucleoplasm by fluorescence resonance energy transfer (FRET) (Chusainow *et al.*, 2005; Rino *et al.*, 2008). A recent study provided a map of SR protein splicing complexes in the nucleus, and showed they act in exon and intron definition *in vivo* by identifying interactions of individual SR proteins with U1-70K at the 5' ss or U2AF35 at the 3' ss (Ellis *et al.*, 2008).

The U12-type class of pre-mRNA introns, also known as AT-AC introns, are spliced by the less abundant U12-dependent (minor) spliceosome. The 5'ss and BP

sequences are highly conserved in AT-AC introns, unlike the degenerate sequences found in GT-AG introns (Will and Luhrmann, 2005). The SR proteins have been shown to participate in AT-AC intron splicing where they promote binding of the U11 and U12 snRNPs to the 5'ss and BP, respectively (Hastings and Krainer, 2001). There is also evidence that SR proteins contact the pre-mRNA of U12-type introns directly via their RS domain, again in an analogous fashion to that seen in conventional splicing (Shen and Green, 2007).

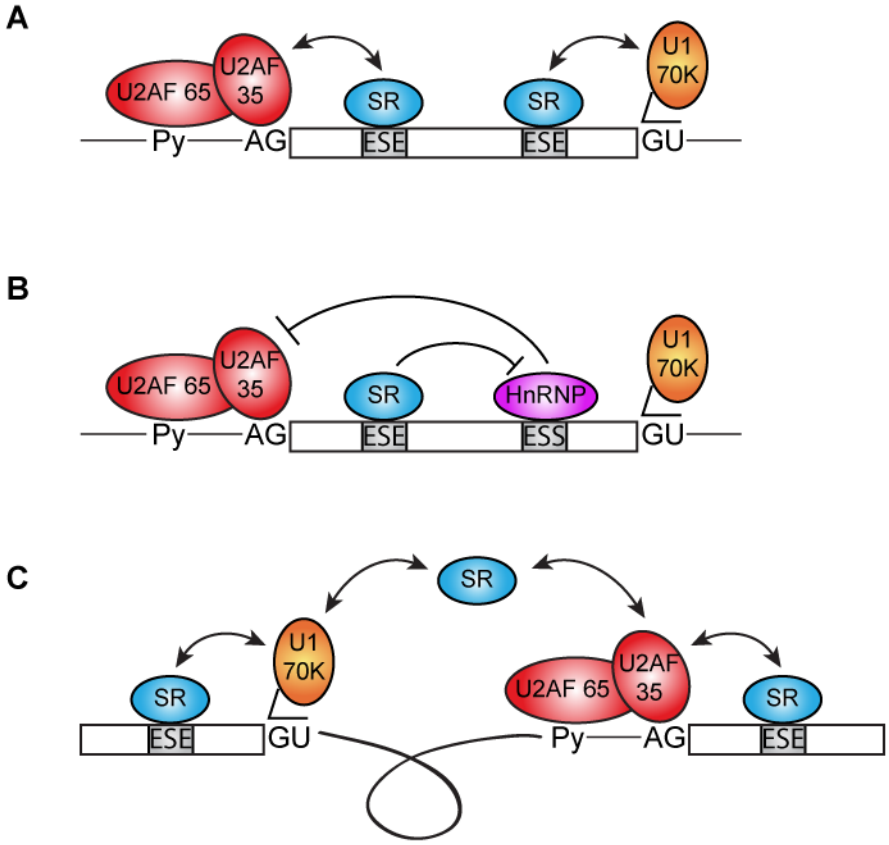


Figure 1.8 Roles of SR proteins in splice site selection (**A**) SR proteins bound to ESE elements recruit U2AF35 to an upstream 3'ss and U1-70K to the downstream 5'ss (**B**) ESSs recruit hnRNP proteins which block 3'ss selection by U2AF. SR proteins bound to ESEs can antagonise the action of these splicing repressors, thereby promoting splice site selection. (**C**) SR proteins can facilitate intron bridging interactions by binding, via the RS domain, to U1-70K and U2AF 35, thereby juxtaposing the 5' and 3'ss.

A delicate interplay of cis-acting sequences and trans-acting factors modulate the splicing of regulated exons in a combinatorial fashion (Smith and Valcarcel, 2000). SR family proteins antagonise the activity of hnRNP A/B proteins in splice site selection, with an excess of hnRNP A1 favouring distal 5'ss, whereas SF2/ASF promotes the use of proximal 5'ss (Mayeda and Krainer, 1992; Eperon *et al.*, 1993; Caceres *et al.*, 1994; Yang *et al.*, 1994; Eperon *et al.*, 2000). Thus, the ratio of

hnRNP A1 to SR proteins in the nucleus is of great importance in alternative splicing regulation and may have a crucial role in the tissue-specific and developmental control of regulated splicing. Accordingly, the protein levels of SF2/ASF and hnRNP A1 have been found to vary naturally over a wide range in rat tissues and also in immortal and transformed cell lines (Zahler *et al.*, 1993b; Hanamura *et al.*, 1998;). SF2/ASF and hnRNP A1 have been found to have an antagonistic role in the regulation of the neuronal-specific N1 exon of the c-src gene (Rooke *et al.*, 2003). Antagonism between hnRNP proteins and SR proteins has also been shown to regulate a highly complex pattern of mutually exclusive exons in the Dscam gene in Drosophila (Olson *et al.*, 2007). A subset of SR proteins has been shown to activate alternative splicing of the cardiac troponin T (cTNT) exon 5 by directly interacting with a purine-rich ESE. Thus, regulation of the levels of individual SR proteins may contribute to the developmental regulation of alternative splicing in cTNT (Ramchatesingh *et al.*, 1995). Interestingly, individual SR proteins can sometimes have antagonistic effects on splice site selection, as is the case with SRp20 and SF2/ASF in the regulation of SRp20 pre-mRNA alternative splicing (Jumaa and Nielsen, 1997) and of SF2/ASF and SC35 in the regulation of β-tropomyosin (Gallego *et al.*, 1997) and human growth hormone pre-mRNA alternative splicing (Solis *et al.*, 2008). Other, non-classical, SR proteins, including p54, SRp38 and SRp86 function solely as negative regulators of alternative splicing, antagonising classical SR proteins and promoting exon skipping (Zhang and Wu, 1996; Cowper *et al.*, 2001; Barnard *et al.*, 2002).

It should also be taken into account the additional importance of structural elements in splice site selection. For example, RNA structure elements associated with alternative splice site selection have been recently identified in the human genome (Shepard and Hertel, 2008). In addition, RNA folding has been shown to affect the recruitment of SR proteins to mouse and human ESE elements in the fibronectin EDA exon (Muro *et al.*, 1999; Buratti *et al.*, 2004). Finally, competing intronic RNA secondary structures help to define a complex pattern of mutually exclusive exons in the Dscam gene (Graveley, 2005).

1.9. Identification of RNA-binding protein targets

Several approaches have been taken to identify physiological RNA targets of RNA-binding proteins. One such approach, termed SELEX (selected evolution of ligands through exponential enrichment) involves the selection of high affinity binding sites from randomised pools of RNA sequences (Tuerk and Gold, 1990). This has resulted in the identification of high-affinity binding sites for SF2/ASF and SC35 (Tacke and Manley, 1995), SRp40 (Tacke *et al.*, 1997), 9G8 and SRp20 (Cavaloc *et al.*, 1999) (Table 1.5). These binding sites consist of purine-rich sequences that resemble 5'ss or exonic sequences, known to function as splicing enhancers. SELEX was also used to identify the consensus high-affinity binding site for hnRNP A1, UAGGGA/U, which also resembles the consensus sequences of vertebrate 5' and 3' splice sites (Burd and Dreyfuss, 1994). SELEX can be used in conjunction with large-scale bioinformatic screens to identify further potential binding sites. An alternative to the SELEX approach, was the development of a functional SELEX strategy, which involves selection for a sequence that will promote splicing, rather than binding alone (Coulter *et al.*, 1997). This can be modified further by performing the splicing reactions in an S100 extract with the addition of a single SR protein (Liu *et al.*, 1998; Liu *et al.*, 2000). The motifs identified for the SR proteins SF2/ASF, SC35, SRp40 and SRp55 using functional SELEX were more redundant than those found by conventional SELEX, suggesting that the specificity of binding *in vivo* depends on other factors than just sequence recognition (Table 1.5). These motifs have been integrated into a web-based program, known as ESE finder, where input sequences can be scanned for potential ESEs responsive to the above proteins (Cartegni *et al.*, 2003).

A number of computational approaches have also been used to define splicing sequence motifs that regulate exon inclusion. RESCUE (Relative Enhancer and Silencer Classification by Unanimous Enrichment)-ESE predicts sequences which could function as ESEs by statistical analysis of exon-intron and splice site composition (Fairbrother *et al.*, 2002). This is based on the observation that ESEs function in a highly position-dependent fashion and are present in constitutively spliced exons and absent in introns. This approach identified 238 candidate ESEs that occurred more frequently in exons with weak splice sites than exons with strong

splice sites. By sequence similarity these were condensed to ten RESCUE-ESE motif clusters and were shown to have enhancer activity *in vivo*. Another computational study found 2,000 RNA octamers, identified as putative exonic splicing enhancers (PESEs) that were found more frequently in exons than in pseudo-exons or intronless genes (Zhang and Chasin, 2004). Validation of a subset of these PESEs resulted in 82% exhibiting decreased splicing efficiency when the PESE was mutated (Zhang *et al.*, 2005).

Table 1.5. RNA sequences identified as SR protein binding sites

SR protein	Binding Site	Characterisation	References
SF2/ASF	RGAAGAAC AGGACRRAGC SRSASGA UGRWG	SELEX SELEX Functional SELEX CLIP	Tacke and Manley, 1995 Tacke and Manley, 1995 Liu *et al.*, 1998 Sanford *et al.*, 2008
SC35	AGSAGAGUA GUUCGAGTUA GRYYNSYR UGUUCSAGWU GWUWCCUGCUA GGGUAUGCUG GAGCAGUAGKS AGGAGAU UGCNGYY	SELEX SELEX Functional SELEX SELEX SELEX SELEX SELEX SELEX Functional SELEX	Tacke and Manley, 1995 Tacke and Manley, 1995 Liu *et al.*, 2000 Cavaloc *et al.*, 1999 Cavaloc *et al.*, 1999 Cavaloc *et al.*, 1999 Cavaloc *et al.*, 1999 Cavaloc *et al.*, 1999 Schaal and Maniatis, 1999
SRp20	GGUCCUCUUC WCWWC CUCKUCY	Gel shift Splicing assay RNA affinity	Lou *et al.*, 1998 Cavaloc *et al.*, 1999 Schaal and Maniatis, 1999
SRp75	GAAGGA	UV crosslinking	Exline *et al.*, 2008
SRp40	GAGCAGUCGGCUC ACDGS	SELEX Functional SELEX	Tacke *et al.*, 1997 Liu *et al.*, 1998
SRp55/B52	USCGKM UCAACCDGGCGAC	Functional SELEX SELEX	Liu *et al.*, 1998 Shi *et al.*, 1997
9G8	UCAACA ACGAGAGAY GGACGACGAG	UV crosslinking SELEX Functional SELEX	Lynch and Maniatis, 1996 Cavaloc *et al.*, 1999 Schaal and Maniatis, 1999
SRp54	C Rich	UV crosslinking and competition	Kennedy *et al.*, 1998
SRp30c	GACGAC AAAGAGCUCGG CUGGAUU	Functional SELEX Functional SELEX Gel shift	Tian and Kole, 2001 Tian and Kole, 2001 Simard and Chabot, 2002
hTra2β	(GAA)n	SELEX	Tacke *et al.*, 1998
SRm160	Purine rich (GAA)n	Splicing assay	Eldridge *et al.*, 1999

N: any nucleotide; R: purine; Y: pyrimidine; S: G or C; K: U or G; W: A or U; D: A, G or U; M: A or C.

Chromatin Immunoprecipitation (ChIP) can be used to study nascent RNA-protein interactions, but a variation of this technique named ribonucleoprotein immuno-precipitation (RIP) provides more information on protein-RNA interactions *in vivo* (Niranjanakumari *et al.*, 2002). RIP involves cross-linking the protein-RNA interactions using formaldehyde, followed by immuno-precipitation of the protein-RNA complexes. After reversing the cross-links the RNA can be amplified by RT-

PCR. This technique relies on random hexamer primers to identify unknown RNAs or can be used in conjunction with microarray technology. Reassociation of RNA-binding proteins after cell lysis can complicate the analysis of these results since observed protein-RNA interactions may not necessarily reflect true *in vivo* interactions (Mili and Steitz, 2004).

An adaptation of the SELEX method described above, known as genomic SELEX, uses real genomic sequences rather than random pools, which allows for identification of authentic protein-binding RNA sequences (Singer *et al.*, 1997; Lorenz *et al.*, 2006).

Recently, a novel technique named CLIP (cross-linking and immuno-precipitation) was developed in order to identify *in vivo* RNA targets (Ule *et al.*, 2003). CLIP involves an *in vivo* photo cross-linking step to capture the protein-RNA interactions, followed by partial RNase digestion to generate RNA tags of approximately 60 nucleotides followed by specific immuno-precipitation of the protein of interest. An advantage of this method is that by using an *in vivo* photo cross-linking step, which induces a covalent protein-nucleic acid bond, these interactions are preserved in an intact cell. CLIP was used to characterise the *in vivo* RNA binding targets of the neuronal-specific splicing factor Nova and has allowed the generation of an RNA map to predict splicing regulation dependent on this protein (Ule *et al.*, 2005; Ule *et al.*, 2006). It has also been used for other RNA binding proteins, including SF2/ASF (Sanford *et al.*, 2008) (Table 1.5), and is used here to identify the *in vivo* RNA targets of Acinus and hnRNP A1.

The identification of SR protein targets and the study of how tissue-specific patterns of splicing change, depending on the complement of SR proteins present, can also be analysed by alternative splicing microarrays (Johnson *et al.*, 2003; Pan *et al.*, 2004; Blanchette *et al.*, 2005).

1.10. SR proteins have functional specificity

Initially, the ability of different individual SR proteins to complement splicing-deficient extracts suggested that SR proteins may have redundant functions. However, the sequence-specific RNA binding ability of individual SR proteins and

differences in their ability to regulate alternative splicing suggested otherwise (Zahler *et al.*, 1993b; Screaton *et al.*, 1995; Wang and Manley, 1995).

A growing body of evidence showed that individual SR proteins were not functionally equivalent in *Drosophila*, *C.elegans* and mouse models. For instance, SF2/ASF was shown to be an essential factor for cell viability in a chicken cell line and its depletion could not be rescued by expression of SC35 or SRp40 indicating a non-redundant function of SF2/ASF (Wang *et al.*, 1996). Other studies have shown that the SR protein B52/SRp55 is essential for *Drosophila* development (Ring and Lis, 1994; Peng and Mount, 1995). B52/SRp55 was shown not to be essential for the splicing of a number of substrates (Hoffman and Lis, 2000), but specific substrates that were mis-spliced in B52 deficient flies were identified (Kim *et al.*, 2003; Rasheva *et al.*, 2006). Furthermore, B52/SRp55 regulates the inclusion of alternative exon 2 in *eyeless*, a master regulator of eye development in *Drosophila*, resulting in the production of a protein isoform that gives rise to a small-eye phenotype. Conversely, the canonical *eyeless* isoform induces eye overgrowth (Fic *et al.*, 2007).

Use of dsRNA interference (RNAi) to inhibit SR protein function during *C.elegans* development revealed that depletion of the orthologue of the mammalian SF2/ASF (CeSF2/ASF) resulted in embryonic lethality, which indicates an essential, non-redundant, role for this gene during nematode development. By contrast, RNAi-mediated depletion of other SR genes resulted in no obvious phenotype, which is indicative of functional redundancy (Kawano *et al.*, 2000; Longman *et al.*, 2000; Longman *et al.*, 2001). The function of SR proteins has also been studied in mice model systems (reviewed by (Moroy and Heyd, 2007)). All SR-null mice for SRp20 (Jumaa *et al.*, 1999), SC35 (Ding *et al.*, 2004; Wang *et al.*, 2001) and SF2/ASF (Xu *et al.*, 2005) show an early embryonic phenotype indicating that SR proteins are not redundant. However, these essential functions appear to be tissue- or developmental stage-specific, as cultured cells from the knockout mice are viable. Therefore it is likely these effects may be due to differential expression profiles of SR proteins during development, although this remains to be determined. The generation of conditional knockouts has allowed further characterisation of SR protein function in different tissue types or at various developmental time points. Tissue specific deletion of SC35 in the thymus of mice results in decreased thymus size and a major

defect in T-cell maturation (Wang *et al.*, 2001), whereas tissue specific ablation of SC35 in the heart has been shown to cause dilated cardiomyopathy (Ding *et al.*, 2004). SF2/ASF has also been shown to have a role in cardiac function; however its function is in the developmental process of postnatal heart remodelling (Xu *et al.*, 2005). Mice knockouts of other splicing factors including Nova, U2AF26, hnRNP U and hnRNP C also result in embryonic lethality or developmental defects, which highlights the importance of splicing for the correct regulation of biological processes such as embryogenesis and tissue maintenance (Moroy and Heyd, 2007).

1.11. Post-splicing activities of SR proteins

SR proteins also function in mRNA processing reactions that occur after splicing including mRNA nuclear export, NMD and translation (Huang and Steitz, 2005). SR proteins display a nuclear localisation pattern and are found to accumulate in splicing speckles (Lamond and Spector, 2003). However, a subset of SR proteins, which includes SF2/ASF, SRp20 and 9G8, shuttle continuously between the nucleus and the cytoplasm (Caceres *et al.*, 1998), reminiscent of what was found for a subset of hnRNP proteins (Dreyfuss *et al.*, 1993). This suggested that the shuttling SR proteins may function in cytoplasmic processes, or be involved in the transport of spliced mRNA. Indeed, SRp20, 9G8 and SF2/ASF function in the nucleocytoplasmic export of mRNA by interacting with the mRNA nuclear export receptor TAP/NFX1 (Huang and Steitz, 2001; Huang *et al.*, 2003), exhibiting a higher affinity when hypophosphorylated (Huang *et al.*, 2004).

SR proteins have also been implicated in regulating the NMD pathway, whereby mRNAs containing PTCs are targeted for degradation. Increased expression of a subset of SR proteins, including SF2/ASF, SC35, SRp40 and SRp55, strongly enhanced NMD (Zhang and Krainer, 2004). Interestingly, this effect does not appear to be dependent on their nucleocytoplasmic shuttling, suggesting a role for SR proteins in enhancing nuclear steps of NMD. A recent study showed that SF2/ASF has the potential to affect the cellular site of NMD, shifting this process to the nuclear compartment before mRNA release from nuclei (Sato *et al.*, 2008).

SF2/ASF controls alternative splicing of pre-mRNAs encoding the kinases MNK2

and S6K1 that are involved in translational regulation. Increased expression of SF2/ASF results in the production of an isoform of MNK2 which promotes MAP kinase-independent eIF4E phosphorylation, and an unusual oncogenic isoform of S6K1, thereby enhancing cap-dependent translation (Karni *et al.*, 2007). SR proteins have also been shown to directly affect translational regulation. SF2/ASF associates with polyribosomes in cytoplasmic extracts and enhances the translation of an ESE-containing luciferase reporter both *in vivo* and *in vitro* (Sanford *et al.*, 2004). This direct effect of SF2/ASF in the regulation of the translation of SF2/ASF-bound mRNA targets is mediated by the recruitment of components of the mTOR signalling pathway resulting in phosphorylation and release of 4E-BP, a competitive inhibitor of cap-dependent translation (Michlewski *et al.*, 2008). The role of mTOR in the activation of S6K1, which phosphorylates eIF4B and S6K1 promoting translation initiation, may also be enhanced by SF2/ASF (Holz *et al.*, 2005) (Figure 1.9). Other SR proteins have also been reported to function in translation. SRp20 has been shown to function in IRES (internal ribosome entry site)-mediated translation of a viral RNA (Bedard *et al.*, 2007), whereas 9G8 plays a role in translation of unspliced mRNA containing a constitutive transport element (CTE) (Swartz *et al.*, 2007).

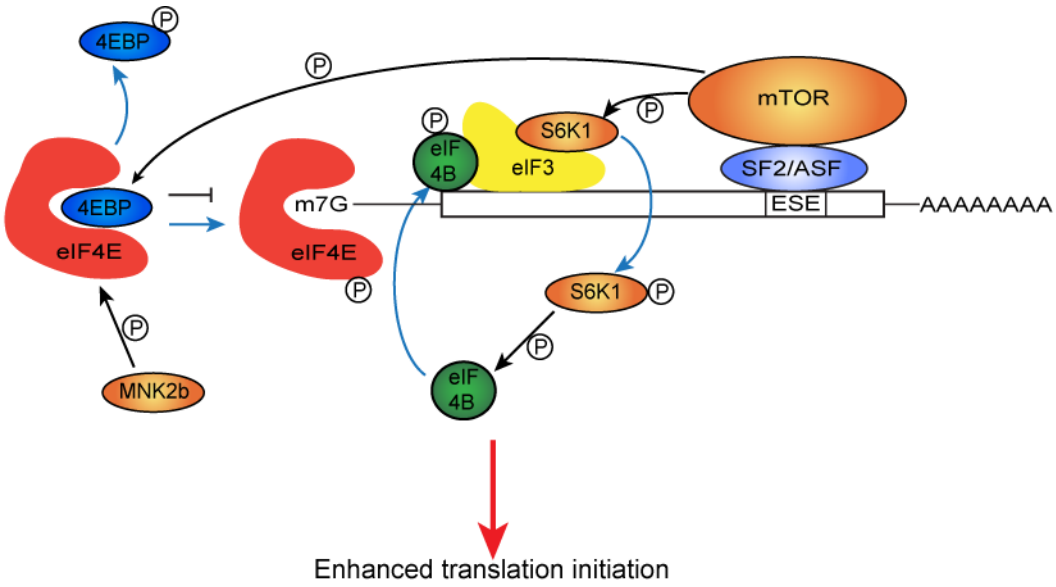


Figure 1.9 Role of SF2/ASF in translation. SF2/ASF-bound mRNAs recruit the mTOR kinase resulting in the phosphorylation and release of 4E-BP, leading to enhanced translation initiation. The mTOR kinase phosphorylates S6K1, which promotes translation initiation, and this may also be enhanced by SF2/ASF. The Mnk2b isoform induced by SF2/ASF-dependent alternative splicing also leads to translation activation.

The results described in this section demonstrate that SR protein function is not restricted to nuclear mRNA splicing, and it seems sensible that proteins already bound to spliced mRNA may function in subsequent processing events as they are already in place to facilitate future interactions. However, it also highlights the requirement for exquisite regulation of SR proteins in order to maintain their role in cytoplasmic processing of mRNAs without disrupting nuclear processes, which are highly sensitive to the relative concentration of splicing factors.

1.12. Splicing regulation by phosphorylation

A dynamic cycle of phosphorylation and dephosphorylation is required for pre-mRNA splicing (Mermoud *et al.*, 1994), this being related, at least in part, to the phosphorylation status of SR proteins. The RS domain of SR proteins is extensively phosphorylated on Ser residues and this plays an important role in regulating the subcellular localisation and activity of SR proteins (reviewed by (Lin and Fu, 2007)). For instance, phosphorylation of the RS domain in SF2/ASF acts to enhance protein-protein interactions with other RS domain-containing splicing factors, such as U1-70K (Xiao and Manley, 1997), whereas dephosphorylation of SR and SR-related proteins is required for splicing catalysis to proceed (Tazi *et al.*, 1993; Cao *et al.*, 1997).

Several protein kinase families have been shown to phosphorylate the RS domain of SR proteins, including the SRPK family (Gui *et al.*, 1994; Wang *et al.*, 1998), the Clk/Sty family dual-specificity kinases (Colwill *et al.*, 1996) and Topoisomerase I (Rossi *et al.*, 1996). SRPK1 is a Ser-specific kinase that binds to a "docking motif" in SF2/ASF which restricts phosphorylation to the N-terminus of the RS domain (Ngo *et al.*, 2008)**.** By contrast, Clk/Sty can phosphorylate the whole of the RS domain resulting in a hyperphosphorylated state (Ngo *et al.*, 2005). The phosphorylation status of the RS domain of SR proteins is also important in the post-splicing activities of SR proteins. A hypophosphorylated RS domain is required for the interaction of nucleocytoplasmic shuttling SR proteins with the TAP/NFX1 nuclear export receptor (Huang *et al.*, 2004). SR protein kinases present in the

cytoplasm are required to re-phosphorylate the RS domain before the SR protein can return to the nucleus (Ding *et al.*, 2006). RS domain dephosphorylation also plays an important role in sorting SR proteins in the nucleus, where shuttling SR proteins and non-shuttling SR proteins are recycled via different pathways (Lin *et al.*, 2005). In the cytoplasm, dephosphorylation of the RS domain enhances mRNA binding of SF2/ASF and contributes to its role in translation (Sanford *et al.*, 2005). SR protein phosphorylation is also important in developmental regulation, as demonstrated in the nematode *Ascaris lumbricoides* (Sanford and Bruzik, 1999).

Importantly, alternative splicing is extensively regulated by signal transduction pathways, whereby signalling cascades can link the splicing machinery to the exterior environment (Lynch, 2007). For instance, the SR protein SRp38 is dephosphorylated upon heat shock by the phosphatase PP1 and becomes a potent splicing repressor (Shin *et al.*, 2004; Shi and Manley, 2007). Two other well described examples are the insulin-induced promotion of PKC beta II alternative splicing as a result of SRp40 phosphorylation by Akt (Patel *et al.*, 2005), and the growth factor induced alternative splicing of the fibronectin EDA exon, via phosphorylation of SF2/ASF and 9G8 by Akt (Blaustein *et al.*, 2004). Interestingly, growth factors not only modify the alternative splicing pattern of the fibronectin gene but also affect its translation in an SR protein-dependent fashion, providing an example where modification of SR protein activity in response to extracellular stimulation leads to a concerted regulation of splicing and translation (Blaustein *et al.*, 2005). Caffeine regulates the alternative splicing of a subset of cancer-associated genes, including the tumor suppressor KLF6. This response is mediated by the SR protein, SC35, which is in turn induced by caffeine, and its overexpression is sufficient to recapitulate this regulated event (Shi *et al.*, 2008).

In addition, the subcellular distribution of hnRNP and SR proteins is regulated by their phosphorylation. Overexpression of Clk/Sty kinases affects the subcellular distribution of SF2/ASF causing its cytoplasmic accumulation (Caceres *et al.*, 1998). Phosphorylation of hnRNP K by ERK leads to its cytoplasmic accumulation (Habelhah *et al.*, 2001), and hnRNP A1 has been shown to be directly phosphorylated by PKA (Cobianchi *et al.*, 1993) and PKC (Municio *et al.*, 1995) resulting in cytoplasmic accumulation of hnRNP A1.

The subcellular localisation of hnRNP A1 can also be modulated by the Mnk1/2 protein kinases that act downstream of p38 in the stress response (van der Houven van Oordt *et al.*, 2000; Guil *et al.*, 2006). Osmotic shock (OSM)-induced phosphorylation was found to negatively regulate the interaction between hnRNP A1 and transportin 1 (Trn1), a transport receptor which mediates hnRNP A1 nuclear import (Allemand *et al.*, 2005), suggesting that the cytoplasmic accumulation of hnRNP A1 is due to impaired nuclear import of hnRNP A1. Stress-induced cytoplasmic accumulation of hnRNP A1 occurs in discrete cytoplasmic foci, the stress granules (SGs), a dynamic cellular component to which stalled translational preinitiation complexes and their associated mRNAs are routed in times of cellular stress (Kedersha and Anderson, 2002). It has also been reported that stress increases the cytoplasmic accumulation of hSlu7, a human splicing factor, and therefore modulates alternative splicing (Shomron *et al.*, 2005).

1.13. Splicing and human disease

The average human gene contains 7 exons of about 145 bp, while the mean intron size is 3,300 bp (Lander et al., 2001). Therefore the mean genomic extent of a single human gene is 27 kb, while the typical coding sequence is only 1,340 bp. This highlights the importance of pre-mRNA splicing in correct gene expression. Furthermore, approximately 15% of mutations that cause genetic disease affect pre-mRNA splicing (Krawczak *et al.*, 1992; Cooper *et al.*, 1997), targeting conserved splicing signals including the 5'ss, 3'ss and BP, as well as enhancer and silencer sequences. Indeed, analysis of a database of 50 single base substitutions associated with exon-skipping in human genes revealed that more than 50% of these mutations disrupted at least one of the target motifs for the SR proteins SF2/ASF, SRp40, SRp55 and SC35 (Valentine, 1998; Liu *et al.*, 2001; reviewed by (Cartegni *et al.*, 2002)).

There is evidence that establishes a connection between the mis-expression of splicing factors and the development of cancerous tissues, mainly as a result of changes in the alternative splicing patterns of key transcripts. Increased expression of SR proteins usually correlates with cancer progression, as shown by elevated

expression of SF2/ASF, SC35 and SRp20 in malignant ovarian tissue (Fischer *et al.*, 2004) and of several classical SR proteins in breast cancer (Stickeler *et al.*, 1999). Altered expression patterns of the hnRNP proteins A1, A2, B1, C1, C2 and K are also observed in lung cancer (Pino *et al.*, 2003; Zech *et al.*, 2006).

SR proteins and hnRNP proteins are implicated in the regulation of genes involved in spinal muscular atrophy (SMA). SMA is a severe hereditary neurodegenerative disorder that results from the lack of a functional SMN1 (survival of motor neuron 1) gene product, which is a key component of the snRNP biogenesis pathway. An SMN1 paralog, the centromeric SMN2 gene, differs by a single nucleotide change, a C>T transition in exon 7 that causes substantial skipping of this exon and results in the production of a non-functional protein. This exon skipping event has been attributed either to the loss of an SF2/ASF-dependent exonic splicing enhancer (Cartegni *et al.*, 2002b) or to the creation of an hnRNP A/B-dependent exonic splicing silencer (Kashima and Manley, 2003; Kashima *et al.*, 2007). Antisense masking of an hnRNP A1/A2 intronic splicing silencer has been shown to correct SMN2 splicing in transgenic mice (Hua *et al.*, 2008). Several therapeutic approaches, which focus on altering the splicing of SMN2 to induce exon 7 inclusion and would result in functional SMN protein in affected patients, have made use of antisense technology (Wirth *et al.*, 2006).

There are many more examples where mis-expression or loss-of-function of splicing factors results in disease. This emphasises not only the necessity for correct splicing regulation but also the therapeutic potential splicing may offer, as detailed above for SMA.

1.14. Splicing and apoptosis

Apoptosis, or programmed cell death, is a genetically controlled process that is fundamental to the development and homeostasis of multicellular organisms. Apoptosis occurs by either the intrinsic or extrinsic pathways, but central to both are caspases (cysteinyl aspartate-specific proteases) (Los *et al.*, 1999) (Figure 1.10). Caspases mediate the effects that give the characteristic morphological changes

associated with apoptosis. These include cell blebbing, cleavage of the nuclear lamins, chromatin condensation and DNA fragmentation.

The extrinsic pathway involves the transmission of extracellular signals to the intracellular death machinery. This occurs by signalling molecules binding to cell surface receptors, such as the Fas and TNFα receptors, and their subsequent activation. The intrinsic pathway is mediated by the mitochondria and can be triggered by oxidative stress, toxins, nitric oxide, high calcium, ceramide and DNA damage. These stresses cause disruption of the outer mitochondrial membrane and the release of proapoptotic factors such as cytochrome c. Both pathways converge onto a common cell-death machinery, which is largely conserved between different organisms.

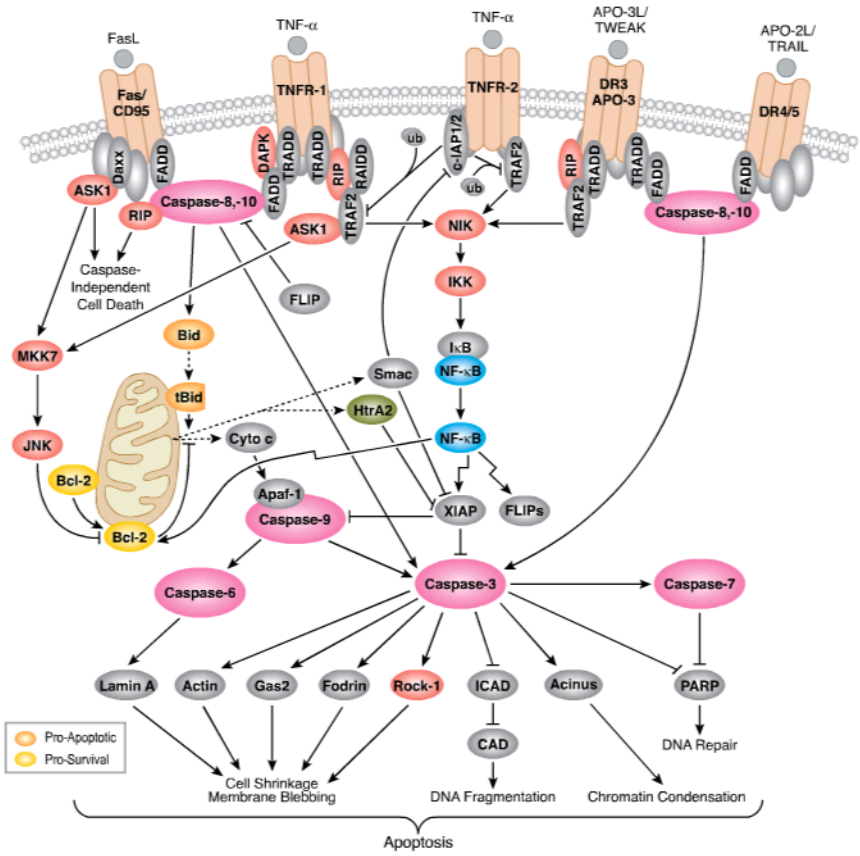


Figure 1.10 Death Receptor Signalling (adapted from www.cellsignal.com). Both the intrinsic and extrinsic apoptotic pathways converge upon caspase-3. This has a number of downstream targets which mediate the morphological changes in cells undergoing apoptosis. Caspases are depicted in pink; kinases in red and transcription factors in blue. Pro-apoptotic factors are shown in orange while pro-survival factors are yellow.

Caspases are maintained as zymogens to prevent unwarranted cell death. Before the 'executioner' caspases (caspase-3, caspase-6 and caspase-7) can attack their cellular substrates their zymogens must be proteolytically cleaved by 'initiator' caspases (caspase-8 and caspase-9) (Earnshaw *et al.*, 1999). Unlike the executioner caspases, each initiator caspase has a long pro-domain and activation is triggered by dimerisation of the zymogen on a dedicated adapter or scaffold protein (Shi, 2002; Boatright and Salvesen, 2003). In the extrinsic pathway the activated death receptors recruit death-domain (DD) containing adapter proteins. These adapter proteins contain further specific domains (e.g., the death effector domain (DED) or the caspase recruitment domain (CARD)) that enable interaction and activation of initiator caspases. In the intrinsic pathway the apoptotic protease-activating factor-1 (Apaf-1) functions in conjunction with cytochrome c to activate caspases via the formation of the apoptosome (Adams, 2003). The apoptosome is a multiprotein caspase-activating complex.

Apaf-1 is regulated by the Bcl-2 family of apoptotic regulators, which are characterised by the presence of up to four conserved α-helical sequences called the Bcl-2 homology (BH) domains. Bcl-2 proteins can be either pro- or anti-apoptotic depending upon which BH domains they contain. The function of Bcl-2 proteins during apoptosis in mammals is complex and involves activation and translocation of proapoptotic members to the outer mitochondrial membrane followed by membrane perturbation and release of apoptogenic factors including cytochrome c (Adams, 2003; Green and Kroemer, 2004).

To complicate matters further caspases also participate in several non-apoptotic cellular processes. Activated caspases exhibit functions during T-cell proliferation and cell cycle regulation, but are also involved in the differentiation of a diverse array of cell types including ethyropoeisis and spermatogenesis (Schwerk and Schulze-Osthoff, 2003). Obviously the apoptotic pathway must be tightly regulated or the consequences to the organism would be catastrophic. The p53 family plays a large role in the transcriptional regulation of the apoptotic program (Melino *et al.*, 2002; Vousden and Lu, 2002).

However, in addition to transcriptional regulation, programmed cell death is also regulated by alternative splicing. For example alternative splicing of Bcl-x, Ced-4

and Ich-1 pre-mRNAs results in the formation of products that have opposite roles in apoptosis (Boise *et al.*, 1993; Wang *et al.*, 1994; Shaham and Horvitz, 1996). In the case of Ich-1 (also known as caspase-2) there are two alternatively spliced forms: Ich-1L, which is pro-apoptotic, and Ich-1S, which is antiapoptotic. SC35 and SF2/ASF promote exon skipping to decrease the ratio of Ich-1S: Ich-1L. In comparison, hnRNP A1 allows exon inclusion to increase the ratio of Ich-1S: Ich-1L. In cultured cells, SC35 overexpression increases apoptosis, while hnRNP A1 overexpression decreases apoptosis (Jiang *et al.*, 1998). Additionally, SC35 has been implicated as a key target of the transcription factor E2F1, which functions in cell cycle progression and apoptosis. E2F1 mediates apoptosis by upregulating SC35 protein expression, thereby switching the alternative splicing profile of apoptotic genes towards the expression of pro-apoptotic splice variants (Merdzhanova *et al.*, 2008).

Many more factors involved in apoptosis are subject to regulation by alternative splicing including death receptors and their ligands, adaptor proteins, members of the Bcl-2 family, caspases and p53 transcription factors (Schwerk and Schulze-Osthoff, 2005). Alternative splicing of a large number of these proteins results in the expression of different isoforms with distinct functional activity, altering their role in apoptosis. However, it is unclear how alternative splicing is used during the apoptotic process. It is unlikely to be an effective mode of regulation during rapid activation of caspase cascades leading to fast cell death, but it may play an important role during slower forms of cell death such as neurodegeneration.

1.15. Acinus

Acinus (apoptotic chromatin condensation inducer in the nucleus) is a protein implicated in apoptosis and RNA splicing. Three isoforms of Acinus are encoded by ACIN1; Acinus L, Acinus S and Acinus S' (Figures 1.11 and 1.12). These isoforms are generated by the use of alternate promoters and alternative splicing (Figure 1.11), resulting in varying N-terminal sequences of Acinus S and Acinus S' which do not align to Acinus L. However, exons 11 – 21 are common to all isoforms, and within this region is a domain homologous to the RRM of the *Drosphila* splicing regulator

sex-lethal (Sxl). In addition to the RRM, all Acinus isoforms contain an arginine-rich region, and unique to Acinus L is a serine-rich region (Figure 1.12). Both these regions contain RS dipeptides similar to that seen in the RS domain of SR proteins, however, only a maximum of five consecutive RS dipeptides are found in either region. The structural and functional properties of the RS domain are important for pre-mRNA splicing and subcellular nuclear localisation, as discussed above. Previous studies have demonstrated a tract of consecutive RS dipeptides are necessary and sufficient to direct nucleocytoplasmic shuttling of specific SR proteins (Cazalla *et al.*, 2002). Therefore, although the arginine-rich region of Acinus is referred to as an RS domain in the literature and hereafter, placing Acinus into the SR-related family of proteins (Table 1.3), it is unclear if the scarcity of RS dipeptides may hinder its functions as a classical RS domain.

ACIN1 is found in vertebrate eukaryotes, but also in fruitflies and the nematode worm. There is a high degree of conservation across species at the protein level (Figure 1.13), although only one isoform could be identified in some species. *D. melanogaster* and *C. elegans* show the most divergence from the conserved protein sequence but despite this the region encompassing the RRM is highly conserved.

Acinus is cleaved during apoptosis by caspase-3 at aspartate 1093 and by an unknown protease at serine 987 (Sahara *et al.*, 1999). This cleaved product contains the RRM (Figure 1.12) and has been reported to cause chromatin condensation (Sahara *et al.*, 1999). Further studies have demonstrated that shRNAi knockdown of Acinus substantially decreases chromatin condensation (Hu *et al.*, 2005), and AMPA-induced excitotoxicity increases the nuclear levels of caspase-activated acinus and chromatin condensation in pyramidal neurons of the hippocampus (Henne *et al.*, 2006). This work suggests that Acinus acts downstream of caspase-3 to mediate chromatin condensation, one of the characteristic morphological changes associated with apoptosis. The molecular mechanism underlying this process is unknown, although various studies have given an insight into how this may be regulated.

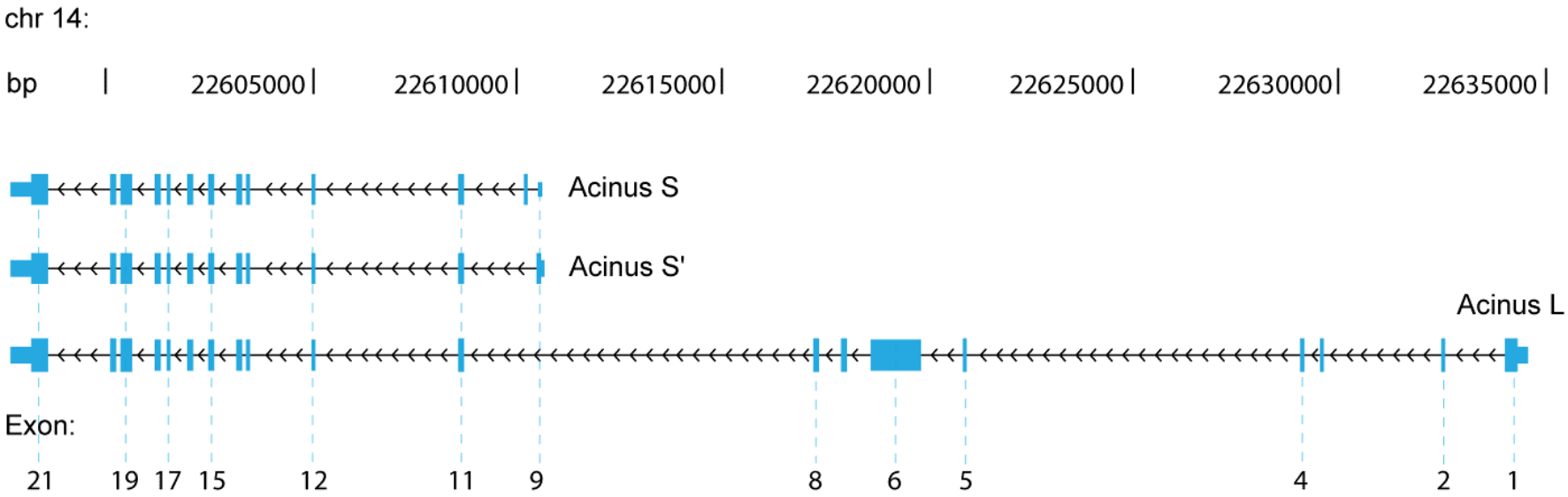


Figure 1.11 ACIN1 is located on chromosome 14 and encodes three isoforms of Acinus generated by use of alternate promoters and alternative splicing. ACIN1 is located on the negative strand of chromosome 14, co-ordinates chr14:22,597,614-22,634,663. Acinus S and Acinus S' are generated from an alternate promoter than that used for Acinus L. Exons 9 and 10 are cassette exons that are skipped in Acinus L. Exon 10 is also skipped in the Acinus S' isoform, while it is included in Acinus S. This results in varying N-terminal sequences of the Acinus S and Acinus S' proteins.

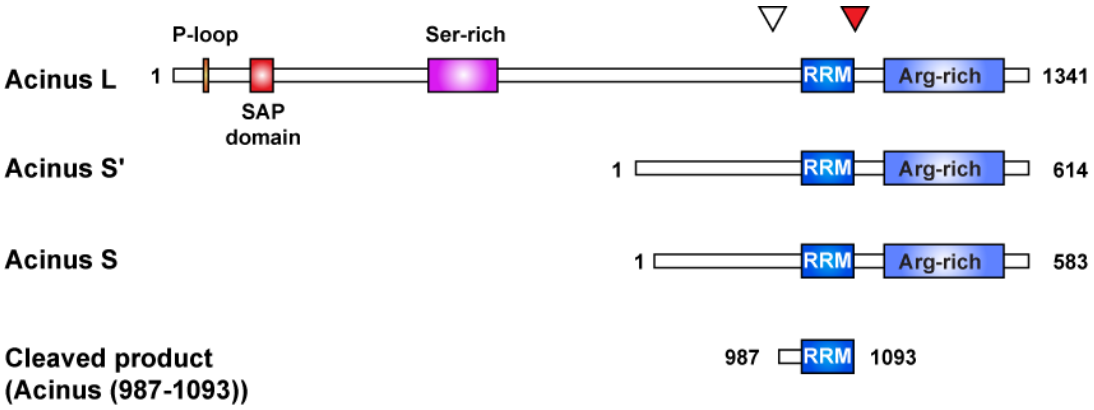


Figure 1.12 Schematic of Acinus isoforms. The apoptotically active form of Acinus is generated by caspase-3 cleavage at aspartate 1093 (red arrowhead) and an unknown protease at serine 987 (white arrowhead). Acinus-L contains a P-loop for nucleotide binding, a SAP domain and a serine-rich region. All isoforms contain a region similar to the RNA recognition motif (RRM) of Drosophila Sxl and an arginine-rich domain at the C-terminus.

Apoptotic chromatin condensation has been linked to the phosphorylation of histone H2B at serine 14. This is mediated by Mst1 (mammalian sterile twenty kinase), which is activated upon cleavage by caspase-3 (Cheung *et al.*, 2003). The caspase-cleaved form of Acinus, but not the full length isoforms, can bind both Mst1 and PKC-δ, resulting in an enhancement of their kinase activities (Hu *et al.*, 2007). Furthermore, Acinus-elicited H2B phosphorylation and chromatin condensation are abrogated in PKC-δ-deficient mouse embryonic fibroblast cells and siRNA knocked down PC12 cells (Hu *et al.*, 2007), suggesting PKC-δ is the kinase downstream of Acinus responsible for chromatin condensation.

Figure 1.13. Acinus is conserved across metazoa. Protein sequences were obtained from NCBI (Refseq accession numbers shown below) and aligned with ClustalW (only the C-terminal regions of the Acinus L isoforms which overlap with the Acinus S and S' isoforms are shown). GeneDoc (Nicholas *et al.*, 1997) was used to show the level of protein conservation between species. Conserved residues are shaded: 100% conservation, black; 80% conservation, dark grey; 60% conservation, light grey. The region showing homology to the RNA recognition motif (RRM) of Sxl is indicated. Hs: Homo sapiens; Pt: Pan troglodytes; Mm: Mus musculus; Xl: Xenopus laevis; Xt: Xenopus tropicalis; Dr: Danio rerio; Dm: Drosophila melanogaster; Ce: Caenorhabditis elegans. HsAcinusL: AAD56724; HsAcinusS: AAD56725; HsAcinusS': AAD56726; PtAcinusL: XP_001160907; PtAcinusS: XP_001160747; PtAcinusS': XP_001160786; MmAcinusL: AAF89661; MmAcinusS: AAD56723; XlAcinus: AAI08815; DrAcinus: AAI16537; DmAcinus: NP_609935; CeAcinus: NP_491344.

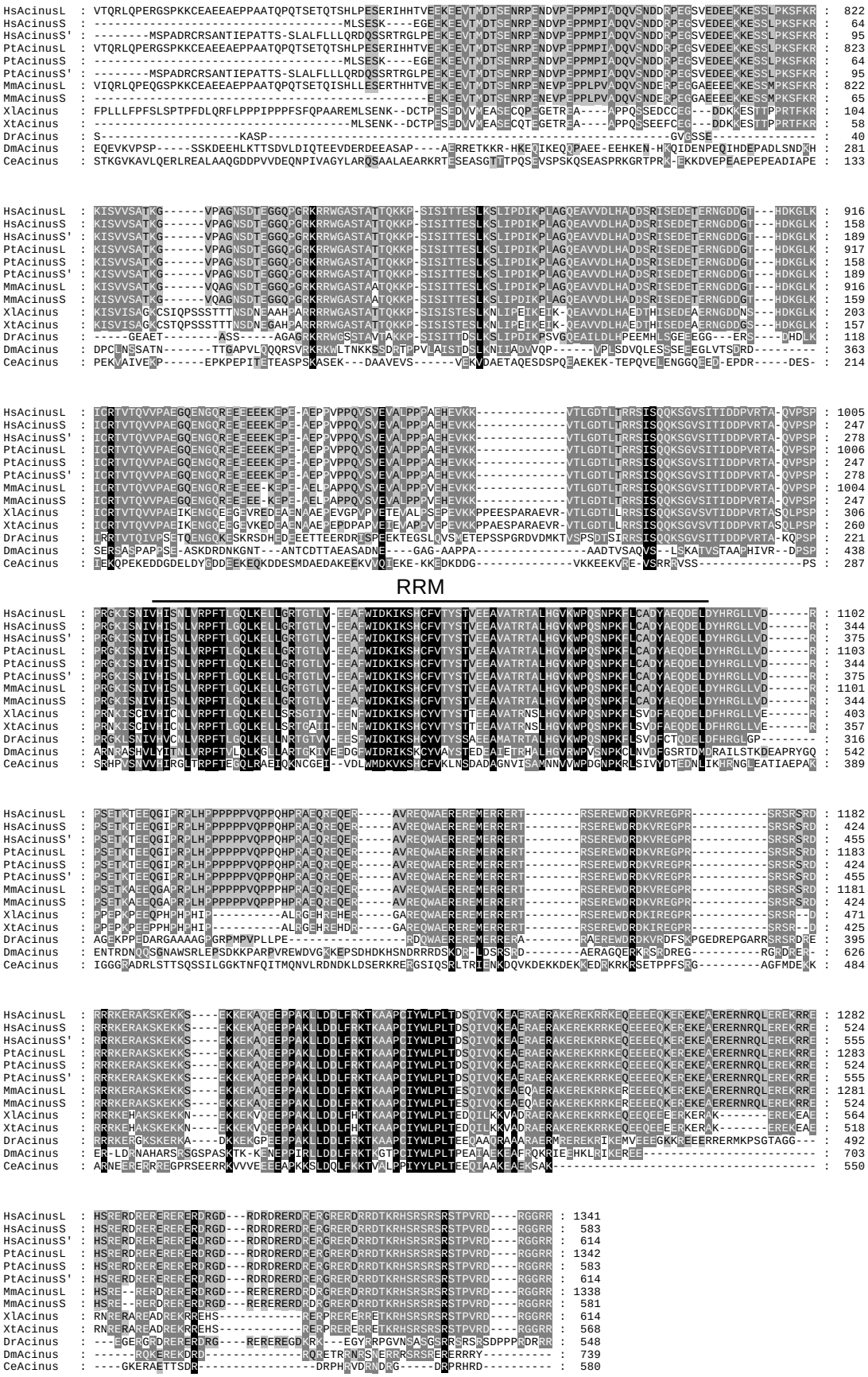


Figure 1.13 Acinus is conserved across metazoa.

Acinus-mediated chromatin condensation is regulated by Akt-dependent phosphorylation. Akt is a kinase involved in numerous signalling pathways, many of these regulating the cell cycle, cell growth or cell death (Franke, 2008). Akt phosphorylates Acinus directly at serines 422 and 573 (Hu *et al.*, 2005), which prevents caspase-3 cleavage of Acinus, thereby inhibiting chromatin condensation. Akt also phosphorylates nuclear zyxin, a focal adhesion molecule, to promote cardiomyocyte survival. The phosphorylation of zyxin induces the formation of a zyxin-acinus complex and results in inhibition of chromatin condensation and antagonism of apoptosis (Chan *et al.*, 2007).

Acinus has also been linked to pre-mRNA splicing. Structurally, it contains an RRM and an RS domain, both of which are known to function in the splicing reaction. The RS domain of Acinus has even been shown to associate directly with nuclear RNA, which may relate to its role in splicing (Nikolakakai *et al.*, 2008). Acinus is one of the numerous proteins to be identified as part of the spliceosome (Rappsilber *et al.*, 2002; Zhou *et al.*, 2002) and is one of the proteins found in the outer core of the EJC (Tange *et al.*, 2005) (Figure 1.4). It is also a member of the ASAP (apoptosis and splicing associated protein) complex, together with SAP18 and RNPS1, which inhibits *in vitro* splicing reactions (Schwerk *et al.*, 2003). Additionally, Acinus is phosphorylated by SPK2, a SR protein-specific kinase that regulates the function of SR proteins in pre-mRNA splicing (Jang *et al.*, 2008). Together these results strongly suggest Acinus acts in the pre-mRNA splicing process, although its precise role remains unknown.

1.16. Project outline

At the outset of this PhD little was known about the function of Acinus. The aims of this PhD were to further characterise the role of Acinus in apoptosis, and to determine if this was linked to its role in RNA processing.

To gain a greater understanding of the role Acinus plays as an RNA-binding protein I set out to identify the *in vivo* RNA targets of Acinus by using the novel

cross-linking and immuno-precipitaion (CLIP) technique. I aimed to optimise this technique for Acinus as part of this PhD.

In addition I used CLIP to identify the targets of another RNA binding protein, hnRNP A1, whose role in RNA processing has been extensively studied. By using CLIP to determine the RNA targets of two RNA-binding proteins I hoped to be able to compare if these two sets of targets are distinct or show some overlap, and if either protein bound specifically to a certain subset of RNAs, for example those encoding proteins involved in RNA processing. Information about the transcripts a protein binds is important for further characterisation of the biological function of the protein.

To summarise, the main aims of this PhD were:

1. Further characterisation of the role of Acinus in apoptosis

2. Optimisation of CLIP for Acinus and identification of the endogenous transcripts to which Acinus binds

3. Identification of the endogenous transcripts to which hnRNP A1 binds, by use of CLIP

Chapter 2: Materials and Methods

2.1. Materials

Table 2.1. Preparation of Materials

Blocking Buffer for Immunoblotting 5% Milk Powder in TBS	**Blocking Buffer for Immunostaining** 1% BSA, 2% serum in PBS
Laemmli Sample Buffer (2X) 50 mM Tris-HCl pH 6.8 100 mM DTT 2% SDS 0.1% Bromophenol blue 10% Glycerol	**Luria Broth (LB)** 10 g NaCl 10 g Bacto-tryptone 5 g Yeast extract Adjust volume to 1 L with dH_2O
LB Agar 10 g NaCl 10 g Bacto-tryptone 5 g Yeast extract 15 g Difco Agar Adjust volume to 1 L with dH_2O	**NuPAGE LDS Sample Buffer (4X)** 106 mM Tris HCl 141 mM Tris base 2% LDS 10% Glycerol 0.51 mM EDTA 0.22 mM SERVA® Blue G250 0.175 mM Phenol Red 50 mM DTT
NuPAGE Running Buffer 104.6 g MOPS 60.6 g Tris Base 10 g SDS 3 g EDTA Adjust volume to 500 ml with dH_2O	**Paraformaldehyde (4%)** 4 g paraformaldehyde in 66ml PBS Heat to 60ºC with stirring (in fumehood) Add 0.5 M NaOH until solution clears Adjust volume to 100 ml with PBS, pH7
PBS 137 mM NaCl 2.7 mM KCl 10 mM Na_2HPO_4 2 mM KH_2PO_4	**PBST** PBS 0.2% Triton X100

TBE (10X) 108 g Tris base 55 g Boric acid 9.3 g EDTA Adjust volume to 1L with dH_2O	**TBS** 20 mM Tris-HCl pH 7.5 137 mM NaCl
TBST 20 mM Tris-HCl pH 7.5 137 mM NaCl 0.1% Tween	**TE** 10 mM Tris-HCl, pH 7.5 1 mM EDTA
Transfer buffer 4.04 g TrisBase 14.4 g Glycine 200 ml Methanol Adjust volume to 1L with dH_2O	**Trypsin** 2 g Trypsin 5 ml Phenol Red (0.2%) 0.06 g Penicillin 0.13 g Streptomycin Adjust volume to 1 L with PBS, pH 7.8
Versene 0.4 g EDTA 5 ml Phenol Red (0.2%) Adjust volume to 1 L with PBS	

2.2. General Molecular Biology Methods

Gel electrophoresis

DNA samples were separated by electrophoresis through a 1% agarose gel made up in 1X TBE buffer supplemented with ethidium bromide at a final concentration of 0.5 μg/ml. Samples were loaded in an appropriate volume of 6X gel loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF and 30% glycerol in water; Sambrook *et al.*, 1989). Sizes of DNA fragments were estimated with the assistance

of either GeneRuler 1 kb DNA ladder (Fermentas) or GeneRuler 100 bp DNA ladder (Fermentas).

Phenol/chloroform extraction and ethanol precipitation

An equal volume of phenol/chloroform (Invitrogen) was added to the sample, followed by centrifugation at 13,000 rpm (Heraeus Biofuge) for 15 minutes. Organic and aqueous components divided into phases and the aqueous, DNA containing, top phase was removed. DNA was precipitated from the aqueous phase by the addition of 3M sodium acetate at $1/10^{th}$ of the sample volume, and 2.5 volumes of ethanol. The DNA was recovered by centrifuging for 15 minutes at 13,000 rpm (Heraeus Biofuge) at 4ºC. The pellet was washed with 70% ethanol to remove residual salt and left to dry at room temperature. The RNA or DNA was re-suspended in an appropriate volume of RNase free water.

PCR purification by gel extraction

PCR products were electrophoresed on an agarose gel and extracted using the QIAGEN gel extraction kit. The manufacturer's instructions were followed, with the additional step of warming the water to 60ºC, which was used to elute the DNA from the filter membrane.

Nucleic acid quantification

DNA and RNA quantification were carried out on a Nano Drop ND-1000 Spectrophotometer. 1.5 µl of sample was loaded onto the pestle, and the absorbance measured using the DNA-50 or RNA-40 settings. Alternatively, RNA was analysed by on-chip gel electrophoresis by the Agilent 2100 Bioanalyzer 2100. The Agilent RNA 6000 Nano Kit was used according to the manufacturer's instructions.

Sequencing

Big Dye v3.1 (Applied Biosystems) sequencing kit was used. The manufacturer's instructions were followed and the sequences read on an ABI Prism® 3730 Genetic Analyser. Sequences were analysed using Sequencher Version 4.0.5 software. ClustalW (http://www.ebi.ac.uk/clustalw/) was used to align DNA sequences.

RNA isolation

At least 5 x 10^6 cells were re-suspended in 1 ml of TRIZOL reagent (Invitrogen) by homogenisation through a 25G needle. RNA was isolated according to the manufacturer's instructions

pGEM-T easy cloning

Taq polymerases frequently add 3' deoxyadenosine to the ends of PCR products in a non-template directed fashion. PGEM-T easy (and other TA vectors) is linearised with 3' deoxythymidine overhangs, this allows efficient ligation of PCR products into the vector cloning site. 2 µl of the gel extracted PCR product was added to the Promega ligation mixture, according to the manufacturer's instructions and incubated at 16ºC overnight. 5 µl of the ligation mixture was transformed into chemically competent *E.coli* DH5α cells (Invitrogen).

Bacterial transformations

5 µl of the ligation mixture was incubated on ice for 10 minutes with 50 µl of chemically competent *E.coli* DH5α cells (Invitrogen). The transformation mixture was then heat shocked at 42ºC for 30 seconds, before being left on ice for a further 5 minutes. 750 µl of antibiotic-free LB was added and the mixture incubated at 37ºC with shaking for 1 hour. The transformed cells were then incubated on agar plates with the appropriate antibiotic. If blue/white selection was employed then X-gal (40 µl at 20 mg/ml) and IPTG (5 µl at 200 mg/ml) were added to the plates before streaking the transformed cells.

Bacterial Cell culture

Cultures were inoculated into, and cells grown in 400 ml LB with the appropriate selective marker (ampicillin at 20 mg/ml, kanamycin at 20mg/ml or chloramphenicol at 13 mg/ml) in an Innova 4200 incubator shaker (New Brunswick Scientific) set to 220 rpm. for >16 hours at 37ºC.

Isolation of plasmid DNA

Plasmids were extracted from the positive transformants with QIAGEN mini-prep kit and QIAGEN maxi-prep kit according to the maufacturer's instructions.

2.3. Cell Culture Methods

Cell Culture

HeLa or 293T cells were grown in DMEM (Invitrogen) supplemented with 10% foetal calf serum (FCS) and 1% penicillin/streptomycin, and incubated at 37ºC in the presence of 5% CO_2. To remove cells from flasks, excess FCS (an inhibitor of trypsin) containing media was washed away with PBS before treatment with versene containing 10% trypsin (Table 2.1).

Cell counting

Cells were harvested and an appropriate dilution made for counting. A 10 µl aliquot of this suspension was spread over an improved neubauer haemocytometer (Hawksley Crystallite BS748; 0.1 mm depth, 1/400 mm^2). The number of cells in four 1 mm^2 areas were counted and the average of these multiplied by 10^4, and appropriately for the dilutions, to determine the number of cells per ml in the suspension. The remaining cell suspension was harvested by centrifugation at 1200 rpm for 5 minutes (WIFUG 500E) before resuspension in an appropriate volume for the experiment.

DNA Transfection

HeLa or 293T cells were grown to 80% confluence in DMEM (Invitrogen) supplemented with 10% FCS, and incubated at 37ºC in the presence of 5% CO_2. Cells were transfected with plasmid DNA by lipofectamine-2000 (Invitrogen) at two concentrations, 2 µg/ml (in accordance with the manufacturer's instructions, and referred to as 'high' in this thesis) and 20 ng/ml (referred to as 'low' in this thesis). The remainder of the transfection procedure was performed in accordance with the manufacturer's instructions.

RNA Transfection

ON-TARGETplus™ siRNA reagents (Dharmacon) were used for knockdown of Acinus and hnRNP A1. A smart pool (a set of 4 siRNA duplexes targeted against the 3'UTR or ORF) was used for depletion of hnRNP A1 (Dharmacon Cat. #: L-008221-00). For Acinus three siRNA duplexes were tested that I designed to be complementary to the C-terminus of the protein, i.e. they target all three Acinus isoforms. The sense sequences are as follows:

siRNA 1 GCAAGAAGAAGAAGAGCAAUU

siRNA 3 CCAAAUCAUUCAAGAGGAAUU

siRNA 5 CUGCAGAGCAUGAAGUAAAUU

Cells were grown to 40% confluency in DMEM (Invitrogen) supplemented with 10% FCS, and incubated at 37ºC in the presence of 5% CO_2. These siRNA duplexes were transfected into HeLa or 293T cells using DharmaFECT®1 (Dharmacon) according to the manufacturer's instructions.

Induction of apoptosis

For apoptotic assays, cells were cultured on glass coverslips and cell death was induced by the addition of either 100 μM etoposide (Sigma), 1 μM staurosporine (Sigma) or 1 μg/μl αFas antibody (MBL) for the time indicated. Cells were stained with DAPI and samples were observed on a Zeiss Axioscop microscope. Images were acquired with a Photometrics CH250 cooled CCD camera using IP Lab software.

Cell viability assay

Cell viability was determined using a 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfopheny)-2H-tetrazolium (MTS) assay (CellTiter 96 Aqueous One Solution cell proliferation assay kit; Promega). In this assay, MTS tetrazolium is converted to a formazan product by metabolically active cells and the quantity of this product can be measured by the amount of 490 nm absorbance, which is directly proportional to the number of living cells. Cells were plated at a density of 15 x 10^4 cells per well of a 6-well plate. The cells were incubated for 48 hours before

addition of the MTS reagent, and absorbance readings were taken 1 hour later using the Nano Drop ND-1000 spectrophotometer.

2.4. Experimental procedures

Plasmid constructs

Sequences encoding Acinus S and the caspase-cleaved Acinus product, Acinus (987-1093), were generated by PCR amplification from the KIAA 0670 cDNA (Ishikawa *et al.*, 1998) and ligated into the pCGT7 expression vector (Caceres *et al.*, 1997). Transcription is driven by the CMV enhancer/promoter, and the coding sequence begins with an N-terminal epitope tag, MASMTGGQQMG, which corresponds to the first eleven residues of the bacteriophage T7 gene 10 capsid protein and is recognised by the T7-tag monoclonal antibody (Novagen). The amplified fragments were designed with XbaI and BamHI restriction sites (sequence recognised by the restriction enzyme is underlined) and subcloned into the XbaI-BamHI sites of pCGT7. The following primers were used:

Acinus (987-1093)

Forward primer 5'-TCGTCTAGATCCGGAGTTTCCATTACCATT-3'

Reverse primer 5'-TCGGGATCCTCAATCCAGCTCATCTTGCTCGGCAT-3'

Acinus S

Forward primer 5'-TCGTCTAGATTATCAGAAAGCAAAGAAGGTGAG-3'

Reverse primer 5'-TTGGGATCCCTAGCGGCGCCCACCCCG-3'

Immunofluorescence Microscopy

Cells were fixed with 4% *p*-formaldehyde in PBS for 15-30 minutes at room temperature, and permeabilised with 0.2% Triton X-100 for 5 minutes. Endogenous proteins were visualised by a 1 hour incubation at room temperature of the following antibodies at the dilutions specified: mouse anti-hnRNP A1 (4B10, ImmuQuest) (1:500), goat anti-TIA-1 (sc-1751, SantaCruz) (1:200), rabbit anti-Acinus (Ab-2, Calbiochem) (1:1000), anti-SC35 (ab11826, AbCam) (1:1000). The cells were washed three times with PBS and incubated for 1 hour at room temperature with 1:500 dilution of FITC-conjugated anti-mouse IgG, Texas Red-conjugated anti-goat IgG, or FITC-conjugated anti-rabbit IgG. After washing three times, the coverslips

were inverted and mounted on glass microscope slides. Immunofluorescence analysis of T7-tagged proteins was carried out using a 1:1000 dilution of anti-T7 monoclonal antibody (Novagen) followed by a 1:200 dilution of Texas red-conjugated anti-mouse IgG or FITC-conjugated anti-mouse IgG. Samples were observed on a Zeiss Axioscop microscope and images acquired with a Photometrics CH250 cooled CCD camera using IP Lab software.

Cell Fractionation

Fractionation was performed as described (Pinol-Roma and Dreyfuss, 1992). All fractionation steps were carried out on ice. Cells grown in monolayer culture were collected with a cell scraper in PBS. Cells were centrifuged at 800g for 3 minutes at 4 ºC. The pellet was resuspended in PBS and spun at maximum speed for 10 seconds. The pellet was then resuspended in RSB-100 (10 mM Tris-HCl, pH 7.5, 100 mM NaCl, 2.5 mM $MgCl_2$, 20 μg/ml digitonin (Calbiochem)). The cells were incubated on ice for 5 minutes and the soluble cytosolic fraction was separated from the nuclear and digitonin-insoluble fractions by centrifugation at 2000g for 8 minutes at 4 ºC. The supernatant (cytoplasmic fraction) was collected and the remaining pellet was resuspended in RSB-100 containing 0.5% Triton X-100 and disrupted by two 5 second exposures to sonication using a microtip sonicator at 30% power. Following incubation on ice for 5 minutes the sonicated material was centrifuged at 4000g for 15 minutes at 4 ºC, and the supernatant (nuclear fraction) collected.

In vivo *UV Cross-Linking followed by poly A^+ affinity selection*

Cells were cross-linked at 400 mJ/cm^2, then fractionated as above, or a total cell lysate made by addition of lysis buffer (20 mM Tris- HCl, pH 7.5, 5 mM $MgCl_2$, 100 mM KCl, 0.3% NP-40, 0.5% Triton X-100). Samples were normalised for RNA content, and an equal volume of 2X binding buffer (1 M NaCl, 20 mM Tris-HCl (pH 7.5), 1% SDS, 0.2 mM EDTA, mini-complete EDTA-free protease inhibitor tablet (Roche)) added to the extract. Sample incubated with oligo dT cellulose (Ambion) for 1 hour at room temperature on a rotating wheel. Cellulose was washed three times in binding buffer (0.5 M NaCl, 10 mM Tris-HCl (pH 7.5), 0.5 % SDS, 0.1 mM

EDTA, mini-complete EDTA-free protease inhibitor tablet (Roche)), and resuspended in elution buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, mini-complete EDTA-free protease inhibitor tablet (Roche)) supplemented with RNase A/T1 cocktail (Ambion) and incubated for 1 hour at 37 ºC in a thermo-mixer (eppendorf). An equal volume of 20% TCA was added to the eluate and the sample precipitated on ice for 10 minutes followed by centrifugation at 13000 rpm for 10 minutes at 4 ºC. The pellet was washed two times in ice-cold 90% acetone, air-dryed and resuspended in 2X SDS-PAGE sample buffer.

Immuno-Precipitation (IP)

Anti-acinus antibody (Ab-2, Calbiochem) was bound to protein A-sepharose beads (GE Healthcare) at 1:1000 dilution. Following three washes with lysis buffer (20 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 100 mM KCl, 0.3% NP-40, 0.5% Triton X-100) the beads were added to the fractionated cell extract and incubated on a rotating wheel for 1 h at room temperature. Beads were then washed three times with lysis buffer and resuspended in 2X SDS-PAGE sample buffer. For IPs of T7-tagged Acinus, T7-tag antibody agarose beads (Novagen) were used.

Western Blot Analysis

Protein samples isolated from HeLa or 293T extracts were resolved by SDS-PAGE. Proteins were then transferred to nitrocellulose membranes (Whatman, S&S). Non-specific binding sites were blocked by incubation of the membrane with 5% non-fat dry milk in TBS (20 mM Tris HCl (pH 7.5), 137 mM NaCl). Proteins were detected using the following primary antibodies diluted in 5% non-fat dry milk in TBS: rabbit polyclonal anti-Acinus (Ab-2, Calbiochem) (1:5,000), rabbit polyclonal anti-Acinus (Ab7355, Abcam) (1:500), mouse monoclonal anti-hnRNP A1 (4B10, ImmuQuest) (1:1000), rabbit polyclonal anti-GAPDH (Ab9485, Abcam) (1:1000) and mouse monoclonal anti-T7 (Novagen) (1:10,000). Following washing in TBST (20 mM Tris HCl (pH 7.5), 137 mM NaCl, 0.2% Tween), blots were incubated with the appropriate secondary antibodies conjugated to horse-radish peroxidase (Pierce) and detected with the Super Signal West Pico detection reagent (Pierce).

Cross Linking and Immuno-Precipitation (CLIP) assay

CLIP was performed as described (Ule *et al.*, 2005) with a few modifications (see Table 3.1 for modifications made for Acinus CLIP). Briefly the procedure is as follows: Bead preparation was performed as described above for the IP but the lysis buffer was replaced with Pxl buffer (1 X PBS, 0.125% SDS, 0.5% Triton X-100, 0.5% NP-40). Cells were crosslinked at 300 mJ/cm^2 and fractionated as described above. Each fraction had final concentrations of 150 mM NaCl, 0.5% Triton X-100, 0.125% SDS and 0.5% NP-40. Samples were DNase treated (RQ1 DNase, Promega) and RNase treated (RNase A/T1 cocktail, Ambion). For each experiment a range of RNase concentrations must be tested, generally three RNase dilutions are used at final concentrations between 1:1,000 and 1:50,000. The treated cell extracts were added to the prepared beads and incubated on a rotating wheel for 1 h at room temperature. The beads were washed twice with Pxl buffer and twice with 1X PNK buffer (50 mM Tris-Cl, pH 7.5, 10 mM $MgCl_2$, 0.5% NP-40). On-bead phosphatase treatment was performed (Calf intestinal alkaline phosphatase, Roche) before washing the beads twice with 1X PNK + EGTA buffer (1X PNK supplemented with 20 mM EGTA) and then twice with 1 X PNK buffer.

3' RNA linkers (5'-P GTG TCA GTC ACT TCC AGC GG 3'-puromycin) were ligated on bead at 16 ºC overnight. The beads were washed three times with 1X PNK buffer before the on-bead kinase labelling reaction with [γ-^{32}P]ATP was performed (PNK enzyme, NEB). Following three washes with 1X PNK buffer the beads were resuspended in NuPAGE loading buffer (Invitrogen). Samples were electrophoresed on a Novex NuPAGE 4-12% Bis-Tris gradient gel (Invitrogen), transferred to nitrocellulose membrane (Whatman, S&S) and exposed to X-ray film. Thin bands were cut from the membrane from the low RNase digested sample at about 20 kDa above the molecular weight of the protein. The nitrocellulose piece was incubated with Proteinase K before Phenol/Chloroform extraction of the RNA and precipitation overnight at –20 ºC. The pellet was washed twice with 75% ethanol and air-dryed. 5' RNA linkers (5'-OH AGG GAG GAC GAU GCG G 3'-OH) were ligated at 16 ºC overnight. Samples were DNase treated (RQ1 DNase, Promega) before the RNA was extracted by Phenol/Chloroform and precipitated overnight at –20 ºC. The pellet was washed twice with 75% ethanol and air-dryed. The RNA was reverse-transcribed

(SuperScript III, Invitrogen) and then PCR amplified. Primers: forward, 5'-AGGGAGGACGATGCGG-3' and reverse, 5'-CCGCTGGAAGTGACTGACAC-3'. The PCR product was run on a 2% agarose gel, the DNA excised and extracted by Phenol/Chloroform and precipitated overnight at –20 ºC. The pellet was washed with 70% ethanol, air-dried and resuspended prior to cloning into the pGem T-Easy vector (Promega), blue-white selection of positive transformants and sequencing of the inserts, or CLIP tags.

Primer design

Primer pairs were designed to detect alternative splicing events in transcripts identified in CLIP. This was assisted by using the Primer 3 web-based program (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and use of the Blencowe Lab's Desktop PCR program, which selects primers to detect exon skipping events. Settings were modified to target the correct sequence and obtain the desired size of DNA fragment. Where possible, primers that had a GC clamp were selected. See Table 2.2 for primer sequences and locations.

Primers used to amplify GAPDH and Beta-Actin:

GAPDH forward:

AGCCACATCGCTCAGACACC

GAPDH reverse:

TCTACATGGCAACTGTGAGG

Beta-Actin forward:

CATGGATGATGATATCGCCG

Beta-Actin reverse:

ATACTCCTGCTTGCTGATCC

Transcript	Forward primer sequence	Forward primer location	Reverse primer sequence	Reverse primer location	Long product (bp)	Skipped product (bp)
GALNS	CCGAATTTGGACCGGATGGC	Exon 2	CGTTTCTGGCATGGGCGTTG	Exon 3	885	151
GANAB	ATCTCCAGAAAGCCGCCTGTC	Exon 23	AGGATGTTCTTTAGCAACCTCACTC	Exon 23	678	571
MYST1	CATCAGTACCCTGCAATCCCTC	Exon 10	CCAAAGTCCAGGGCCACCAG	Exon 11	598	304
CACHD1(a)	GATGGGAACTTTAATACCAATGTGTC	Exon 6	TACTGCGGTGTTCGTAGCTGC	Exon 8	260	229
CACHD1 (b)	TCTTGCAGACAACCTGAAATCCAAC	Exon 8	GGAATCCCACTGCGTGCTGG	Exon 10	434	289
NBPF14	GGTGGGATCAAGTGAAAAAGG	Exon 7	CAACAGCCAAGCCAACACG	Exon 13	940	210
RCL1	CAAACCACGTGGACGCGTCT	Exon 1	AGTTCAAATGATTCACCATCAATCCC	Exon 4	604	356
JMJD1C	TACTGGTTATGTAAGCGACCTCG	Exon 3	ACTCTGTATCCATTTAAGGAATAAGGAC	Exon 4	686	466
SHROOM3	CCCTTAGAACATGACTTGCTGTC	Exon 7	CCCTCGGGATCCTCACTGTC	Exon 9	361	287
SMAD2	GCGGAGAAGCAGCTCGCCA	Exon 1	AGTGGTGATGGCTTTCTCAAGCT	Exon 3	494	299
FER	TTTCGCTAGACTTCACTGCACGT	Exon 3	GTTTCCAGTAACCGTAATTCCCAG	Exon 5	348	255
KIF1B	CAAAAAGGAGAAGGAAGAAGCA	Exon 24	CAGCCAGAGATCGGGTTTCA	Exon 26	3762	167
IL6R	GTCAAGGACCTCCAGCATCAC	Exon 6	GGTGGACACCTCGTTCTCAG	Exon 8	830	180
SLC25A43	GTGCTCTCCCGTTCTCTGCT	Exon 3	TTGGCTGTCAATCCATTCCAGAG	Exon 5	471	298
XPO6	GCAGTGGCCGTGTGGGTTG	Exon 1	CTGTCATCAGACTTTCCAATGCC	Exon 3	685	549
TMEM189/ UBE2V1	TGGTCGCACACGTACTTTGG	Exon 5	CTCCTTTCTGGCCTTCTTCG	Exon 9	2171	190
BASP1	GGGAGGAGCCCAGCAAGGA	Exon 2	GAGAGAGAGAGGAGAGATAGGAG	Exon 2	396	366
MIRH1	TACACGTGGACCTAACTGCACC	Exon 2	CAGAAGGCAATCATAACCAACCATC	Exon 3	698	313
HNRNPA1 (a)	CTTTGGTGGCAGCCGTGGTG	Exon 7	GGGCCAGAGCTTCTGCCTC	Exon 9	339	183
HNRNPA1 (b)	AAAGCCCTGTCAAAGCAAGAGATG	Exon 6	GGGCCAGAGCTTCTGCCTC	Exon 9	416	185
HMGCR	CAATGCTAAGCATATCCCAGCCT	Exon 12	CTGCTAGTGCTGTCAAATGCCTC	Exon 14	468	309
HN1	CAATGCCGGCGGGTTAGGG	Exon 1	ATCAAAACCTAATGAAAAATTGGATCCAC	Exon 4	383	288
LPHN2 (a)	GTGCTTCAGACACTCATACTGCT	Exon 19	TGGTACAGAAGGGAGTGAGTCC	Exon 23	700	497
LPHN2 (b)	TCTCAGGTGACATCAATAGCACTTC	Exon 20	TCTTCCATGTCAGGGCTGCTC	Exon 23	695	589
SULT4A1	GACCTGGTGACGATGGTGGA	Exon 6	ACAGGCTTTATCCTTACGGTCCA	Exon 8	650	523
FAM38A	GGCTGGTGCCGTTCCTGGT	Exon 33	CCAGCTTCAGGGTGACGGTG	Exon 35	400	321
CRKL	AGGGCTTTTCCCCTTTACGC	Exon 3	AAGAAGCAAGGCCACTGTTCC	Exon 4	3618	340
RPS16	GCGACAGCTGTGGCGCACT	Exon 3	ATCTGGGCCACGTGACCACC	Exon 4	188	137
ARID4A	AGGCCAAGTGATGGAAATAGTGG	Exon 21	TTCTCCTCCTGTCTATGGTTGCA	Exon 23	424	217
CDK4	GCTGATGGACGTCTGTGCCA	Exon 3	TCTCTGCAAAGATACAGCCAACAC	Exon 5	401	233

Table 2.2. Primer sequences used in RT-PCR alternative splicing assays

Statistical tests

Student's t-test and Chi-square analyses were performed using web-based programs (http://faculty.vassar.edu/lowry/newcs.html, http://faculty.vassar.edu/lowry/t_ind_stats.html)

2.5. Bioinformatics

Characterisation of the CLIP tags

All human exon and gene information was downloaded from Ensembl, using the BioMart tool. The sequence data was downloaded using Perl scripts to access the ensembl database through the Ensembl Core Perl API. The CLIP sequences were mapped against unspliced gene, cDNA and coding sequences using Mega BLAST (Zhang *et al.*, 2000). Mapping of the CLIP tags allowed a match to Ensembl gene ids. The characterization of the splicing events in regions in or adjacent to Acinus or hnRNP A1 binding sites was performed according to Ensembl exon information and the FastDB database.

Identification of the hnRNP A1 binding site

The distinct CLIP tags obtained from nuclear or cytoplasmic extract preparations were used on the MEME (Bailey and Elkan, 1994) web server (http://meme.sdsc.edu/meme/meme.html) to identify patterns that are probabilistically significant. The intronic and exonic sets of CLIP tags were used in the same way but did not result in any meaningful patterns.

Modelling CLIP in silico

To generate the random dataset used in Figure 5.9, which gives the distribution of binding sites, in relation to gene structure, that would be expected if an RNA binding protein bound randomly to the transcriptome, an approximation of CLIP was modelled *in silico*. Perl scripts were used to generate a structural transcriptome map, i.e. the co-ordinates of the UTRs, exons and introns of all known coding transcripts. Transcript information was obtained from the March 2006 human UCSC genome

release (http://genome.ucsc.edu). Another perl script was used to model the location of 100 randomly-generated CLIP-like tags of length 60 nt, and repeated 1000 times in order to generate the average number of times tags were present in UTRs, exons or introns.

Chapter 3: Acinus Characterisation

3.1. Introduction

Acinus was identified as a factor acting in the apoptotic pathway that induces chromatin condensation after caspase-3 cleavage (Sahara *et al.*, 1999). Apoptosis, or programmed cell death, is a physiological process whereby unwanted or useless cells are eliminated during development and other normal biological processes, and has an important role in the prevention of a number of diseases including cancer (Thompson, 1995). This is in contrast to necrosis, a form of cell death which occurs when cells are exposed to a serious physical or chemical insult.

Apoptosis involves a series of biochemical events that lead to a variety of morphological changes (Figure 3.1). Early changes involve chromatin condensation along the perimeter of the nucleus. Membrane blebs appear, and the nucleus condenses completely and segregates into several fragments. The cell disintegrates into apoptotic bodies containing the cellular material. Finally, the apoptotic bodies are engulfed by other cells and digested via the lysosomal pathway.

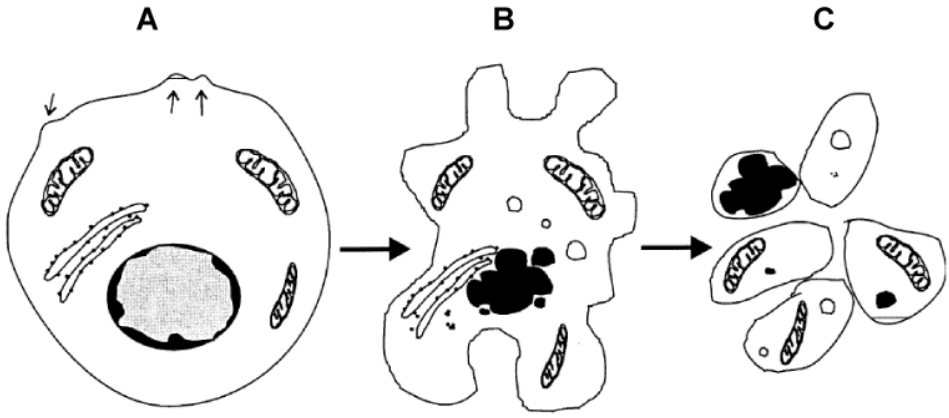


Figure 3.1. Morphological changes associated with apoptosis (adapted from Hacker, 2000). (**A**) Early changes involve condensation of nuclear chromatin along the nuclear periphery and membrane blebbing (arrows). (**B**) The nucleus condenses completely and segregates into several fragments. (**C**) The cell disintegrates into apoptotic bodies containing cellular material. These are taken up by other cells and digested via a lysosomal pathway.

Chromatin condensation is one of the nuclear hallmarks of the apoptotic process. It serves two purposes; it facilitates cell shrinkage, enabling heterophagic removal of the apoptotic cells, and it prevents the DNA of the dead cell, which may include viral genomes or mutated genes, from being incorporated into adjacent cells. Caspases,

and their downstream effectors such as caspase-activated DNase (CAD) and Acinus, are considered to be rate limiting for the development of chromatin condensation (Liu *et al.*, 1997; Enari *et al.*, 1998; Sakahira *et al.*, 1998; Samejima *et al.*, 1998; Sahara *et al.*, 1999). However, caspases and CAD are not the only factors involved in nuclear apoptosis as chromatin condensation has been observed in cells even when caspase activation is inhibited (Hirsch *et al.*, 1997; Deas *et al.*, 1998; Quignon *et al.*, 1998). Apoptosis-inducing factor (AIF), has been identified as one of these additional factors. AIF is a mitochondrial intermembrane flavoprotein, which translocates from mitochondria to nuclei in a caspase-independent fashion where it causes DNA fragmentation and peripheral chromatin condensation (Susin *et al.*, 1999; Daugas *et al.*, 2000). This highlights the complexity of the pathways involved in mediating nuclear apoptosis.

In addition to Acinus' reported role in apoptosis, it has also been identified as a component of the spliceosome (Rappsilber *et al.*, 2002; Zhou *et al.*, 2002) and the EJC (Tange *et al.*, 2005). Structurally, Acinus contains an RS domain and an RRM (Figure 1.12), motifs common in pre-mRNA splicing factors. Interestingly, the cleaved Acinus fragment, shown to be active in mediating chromatin condensation, contains the intact RRM. This raises the question if RNA binding, or the possible role of Acinus in pre-mRNA splicing, is involved in the apoptotic pathway. In order to answer this question I attempted further characterisation of Acinus and its biological functions.

3.2. Results

Acinus is a nuclear protein

Three isoforms of Acinus are produced (Figure 1.12) which can be detected in HeLa cells (Figure 3.2.A) and 293T cells (data not shown) by Western blotting with an antibody that recognises the C-terminal region of all three isoforms. HeLa and 293T cells were used as both are human cell lines and widely available.

Acinus S' and Acinus S have similar molecular weights as a result of only slight variations at the N-terminus (Figure 1.13), resulting in difficulties resolving the two bands.

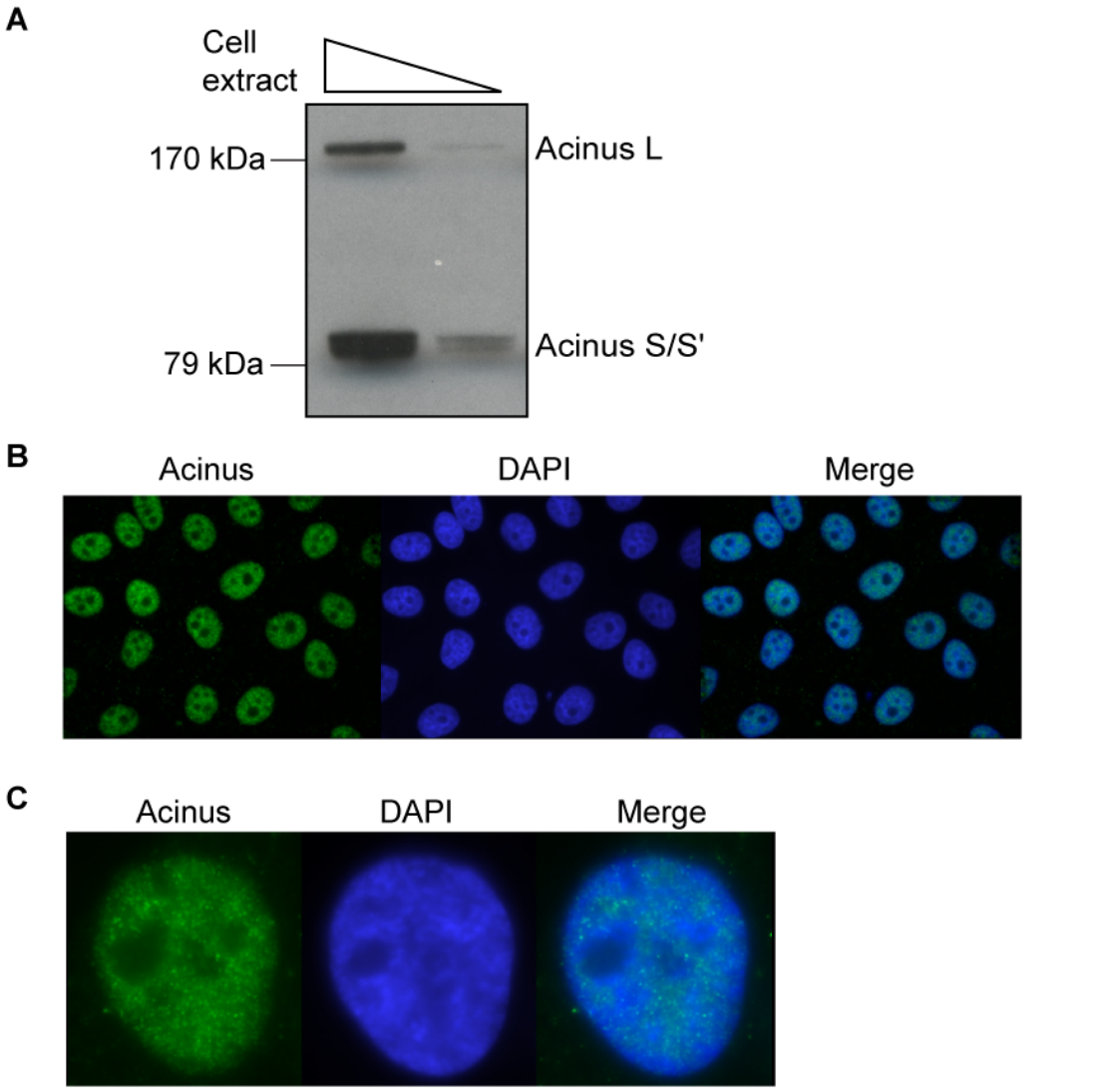


Figure 3.2. Acinus is localised to the nucleus. (**A**) Western blot of HeLa cell extract with an anti-Acinus antibody (Ab-2, Calbiochem), that recognises the C-terminal region of all three Acinus isoforms. (**B**) HeLa cells were immuno-stained with an anti-Acinus antibody (Ab-2, Calbiochem), followed by an anti-rabbit antibody conjugated to FITC. (**C**) Higher magnification image showing the diffuse speckled pattern of Acinus in the nucleus.

The amino acid sequences of Acinus L, Acinus S' and Acinus S (UniProt identifiers: Q9UKV3-1, Q9UKV3-2 and Q9UKV3-3) predict proteins of 1341, 614 and 583 amino acids with calculated molecular masses of 152 kDa, 71 kDa and 68 kDa, respectively. However, the isoforms migrate on SDS-PAGE with apparent

higher molecular masses (Figure 3.2.A). This is generally observed with basic proteins, as they have a high mass to charge ratio resulting in slower migration on SDS-PAGE, but the Acinus isoforms all have isoelectric points around neutral pH. Potential phosphorylation of serine residues within the RS domain could account for the observed anomalous migrations, as has been reported for other RS domain containing proteins including SRrp86, U1-70k, SRp75 and SRrp53 (Barnard and Patton, 2000; Chaudhary *et al.*, 1991; Zahler *et al.*, 1993a; Cazalla *et al.*, 2005).

The subcellular distribution of the endogenous proteins was analysed by indirect immunofluorescence in HeLa (Figure 3.2.B) and 293T cells (data not shown) showing Acinus is exclusively localised to the nucleus. Acinus displays a nuclear diffuse speckled staining, but is excluded from the nucleolus (Figure 3.2.C). Co-localisation with SC35, a nuclear speckle marker, shows Acinus localisation does overlap with SC35 localisation (Figure 3.3) suggesting Acinus is localised to nuclear speckles. However, Acinus localisation is also detected outwith nuclear speckles throughout the nucleoplasm (Figure 3.3). Splicing factors are known to concentrate in nuclear speckles and are recruited from these sites to nascent sites of RNAP II transcription (Misteli *et al.*, 1997).

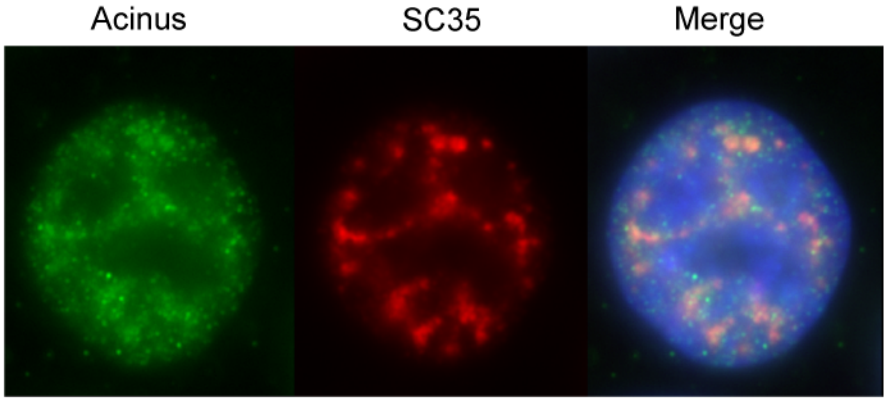


Figure 3.3. Acinus localises to nuclear speckles. SC35 was used as a nuclear speckle marker to show Acinus co-localises to sites of nuclear speckles. HeLa cells were immuno-stained with an anti-acinus antibody (Ab-2, Calbiochem) followed by an anti-rabbit antibody conjugated to FITC, and an anti-SC35 antibody (ab11826, AbCam) followed by an anti-mouse antibody conjugated to TexasRed.

In order to perform over-expression studies, Acinus S was cloned into the pCGT7 vector to produce an N-terminal T7 epitope tagged protein (T7-Acinus S). Cells

transfected with this construct showed the transiently expressed tagged protein also localised to the nucleus (Figure 3.4.A). Furthermore, over-expression of Acinus S appears to lead to an accumulation of the protein in nuclear speckles (Figure 3.4.B). However, a proportion of cells over-expressing Acinus S display a loss of nuclear localisation (Figure 3.4.C). DAPI staining of these cells clearly shows major chromatin condensation has occurred (Figure 3.4.C), suggesting the cells are undergoing apoptosis. During apoptosis the nuclear membrane breaks down, and this is the likely reason for the observed loss of Acinus nuclear localisation.

In order to determine if the biological functions of Acinus were concentration dependent, cells were transfected with T7-Acinus S at two concentrations; 2 µg/ml (referred to as high) and at a concentration 100 fold lower than this, 20 ng/ml (referred to as low). Analysis of Acinus protein expression from cells over-expressing T7-Acinus S at low and high levels, compared to non-transfected or mock-transfected cells, shows transfection at these two concentrations does result in protein expression at two different levels (Figure 3.4.D).

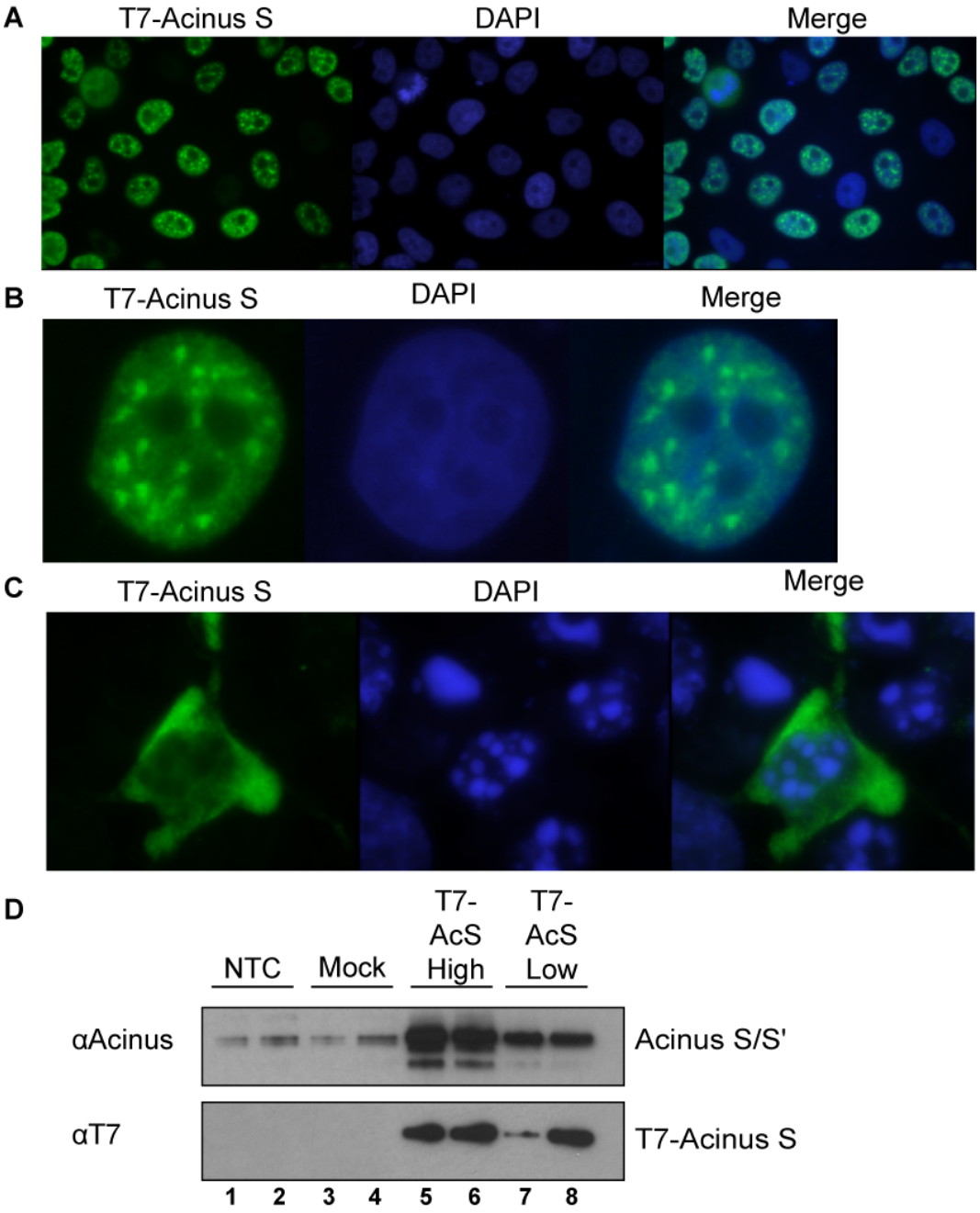


Figure 3.4. Acinus displays a nuclear localisation when over-expressed. (**A**) HeLa cells transfected with a T7-tagged Acinus S construct were immuno-stained with an anti-T7 monoclonal antibody (Novagen), followed by an anti-mouse antibody conjugated to FITC. (**B**) Higher magnification image showing the enlargement of speckle-like bodies when Acinus S is overexpressed. (**C**) Approximately 10% of cells over-expressing Acinus S at low levels and 35% of cells over-expressing Acinus S at high levels undergo apoptosis, compared to 4% of mock-treated cells (see Figure 3.10), which can result in loss of nuclear localisation, most likely due to breakdown of the nuclear envelope. (**D**) Western blot of HeLa cell extracts which had been non-transfected (lanes 1 and 2), mock transfected (lanes 3 and 4), transfected with 2 µg/ml of T7-Acinus S (lanes 5 and 6) or transfected with 20 ng/ml of T7-Acinus S (lanes 7 and 8). Extracts were blotted with an anti-Acinus antibody (Ab-2, Calbiochem) (upper panel) and an anti-T7 antibody (Novagen) (lower panel). NTC: Non-transfected control; AcS: Acinus S.

Expression analysis of Acinus

The expression of splicing factors is regulated in a tissue and developmental-specific manner, as changes in their levels can affect alternative splicing events. For example, the ratio of the antagonistic splicing factors, hnRNP A1 and SF2/ASF, is important for the correct regulation of alternative splicing. Accordingly, the levels of SF2/ASF and hnRNP A1 have been found to vary naturally over a wide range in rat tissues and also in immortal and transformed cell lines (Zahler *et al.*, 1993b; Hanamura *et al.*, 1998). Apoptosis is also regulated in a tissue and developmental-specific manner.

To investigate if the expression of Acinus is tissue-restricted, which may give an insight into the role of Acinus in the splicing or apoptotic processes, RT-PCR analysis was performed to determine the expression of Acinus L, Acinus S' and Acinus S at the RNA level (Figure 3.5). Primers were designed to amplify unique regions of the three isoforms so their expression could be compared. 500 ng of tissue-specific, quality tested RNA purchased from Ambion was used in a one-step RT-PCR reaction. All three Acinus isoforms are widely expressed across a variety of adult human tissues (Figure 3.5). There is some variation in the levels of expression between isoforms, for example Acinus S is only weakly expressed in brain, while Acinus S' is expressed at a much higher level (Figure 3.5.B, lane 3). However, all three isoforms appear to be widely expressed at the RNA level, and do not show tissue restriction amongst the 17 tissues analysed.

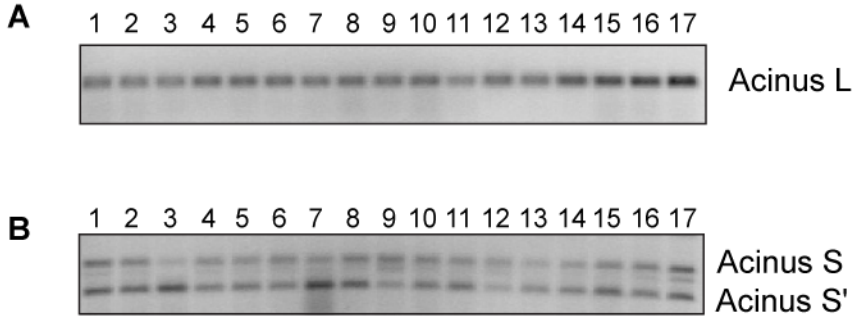


Figure 3.5. Acinus is widely expressed. Acinus L, S and S' were amplified by RT-PCR (SuperScript™ III one-step RT-PCR system with Platinum® Taq, Invitrogen) from a range of tissue-specific RNAs (FirstChoice® Human Total RNA Survey Panel, Ambion).
1: Adipose; 2: Bladder; 3: Brain; 4: Cervix; 5: Colon; 6: Oesophagus; 7: Heart; 8: Kidney; 9: Liver; 10: Lung; 11: Ovary; 12: Placenta; 13: Prostrate; 14: Small intestine; 15: Trachea; 16: Thymus; 17: Thyroid. (**A**) RT-PCR performed with primers specific for Acinus L (Ta 60°C, 40 cycles). (**B**) RT-PCR performed with primers to amplify both Acinus S and S' (Ta 60°C, 40 cycles).

This analysis does not exclude the possibility that Acinus may be expressed at different levels throughout development, or it may be differentially expressed at the protein level which was not tested here. The wide-ranging expression of Acinus may indicate a functional role in cell maintenance. For example, apoptosis is important in maintaining homeostasis in multicellular organisms; the level of cell proliferation must be matched to the level of cell death.

Acinus is an RNA-binding protein

All three Acinus isoforms contain a region that is homologous to the RRM of the *Drosophila* splicing regulator Sxl (Figures 1.12, 3.6.A). To determine if Acinus can bind RNA I used *in situ* UV crosslinking and denaturing oligo dT selection to purify mRNP particles from whole cell extracts. The RNA was then degraded and the associated proteins precipitated (Figure 3.6.B). UV irradiation induces a covalent bond between proteins and RNA, allowing purification of RNA-associated proteins. Therefore proteins should not be precipitated following oligo dT selection of non-UV crosslinked extracts (Figure 3.6.B, lane 3). Equivalent amounts of extracts were blotted before oligo dT selection to control for the expression of proteins in the prepared extract (Figure 3.6.B, Inputs: lanes 1 and 2). Affinity selection of mRNPs by oligo dT cellulose clearly shows that all three isoforms of endogenous Acinus and T7-Acinus S are bound to mRNA (Figure 3.6.B).

SF2/ASF, which has been characterised extensively as an RNA binding protein (Sanford *et al.*, 2005), was used as a positive control in this assay. T7-hnRNP A1 (RRM1,2), a mutated version of hnRNP A1 deficient in RNA binding, was used as a negative control. This protein has two conserved Phe residues in the RNP-1 submotif of each of the two RRMs, which are essential for specific RNA-protein interactions, replaced by Asp residues (Mayeda *et al.*, 1994; Guil *et al.*, 2006).

This result, together with the localisation of Acinus to nuclear speckles (Figure 3.3), suggests Acinus is a factor involved in RNA processing, possibly in pre-mRNA splicing.

A

```
Sxl         LYVTNLPRTITDDQLDTIFGKYGSIVQKNILRDKLTGRPRGVAFVRYN   258
Acinus L    VHISNLVRPFTLGQLKELLGRTGTLVEEAFWIDKIKSH----CFVTYS  1057
            :::::** *.:* .**. ::*: *::*:: :  **:..:    .** *.

Sxl         KREEAQEAISALNNVIPEGGSQPLSVRLAEEHGKAKAAHFMSQMGVV    305
Acinus L    TVEEAVATRTALH-----------GVKWPQSNPKFLCADYAEQDELD   1093
            . ***  : :**:           .*: .:.: *  .*.: .*  :
```

B

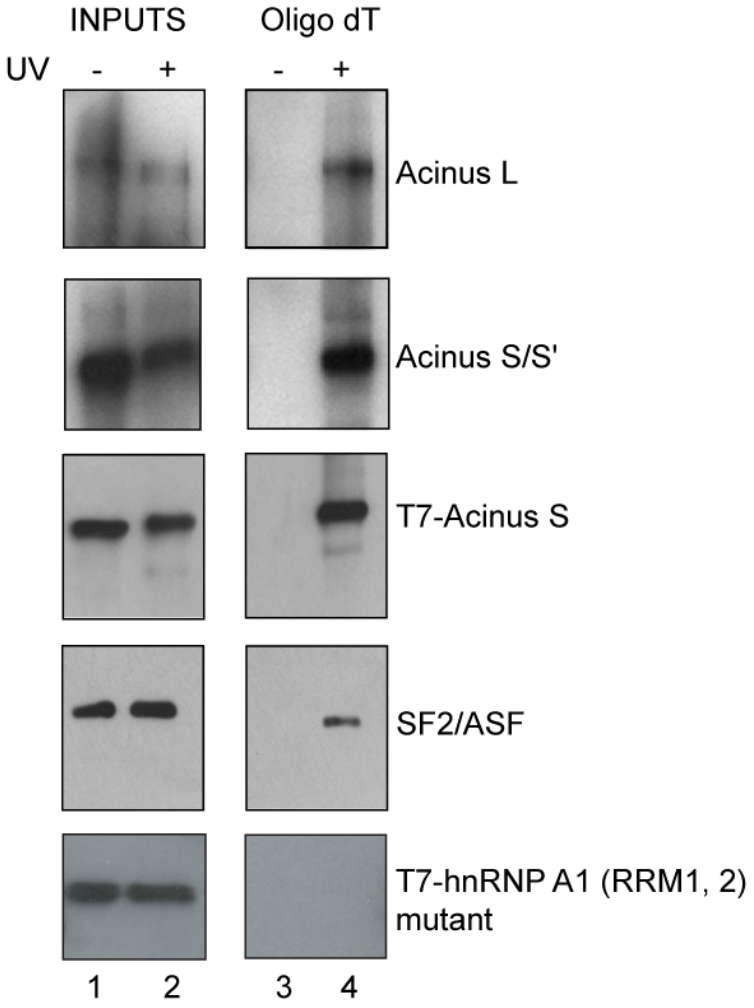


Figure 3.6. Acinus binds RNA. (**A**) Comparison of the amino-acid sequences of Acinus (UniProt accession number, Q9UKV3) and Drosphila Sxl (UniProt accession number, P19339). Residues that are identical between the two proteins are represented by asterisks, while dots indicate positions at which amino acids are functionally similar. (**B**) 293T cells were UV crosslinked, lysed and a nuclear extract prepared. The extract was incubated with oligo(dT) cellulose (Ambion) for 1 h, then eluted for 1 h and RNase digested (RNase A/T1 cocktail, Ambion). Lanes 1 and 2 (Input) contain extracts before oligo dT selection of mRNPs, while lanes 3 and 4 (Oligo dT) contain purified mRNPs eluted from oligo(dT) cellulose. Proteins were precipitated and Western blotted using an anti-Acinus antibody (Ab-2, Calbiochem), an anti-T7 antibody (Novagen) or an SF2/ASF antibody (mAb96, Hanamura *et al.*, 1996).

Classification of apoptotic stages

Extrinsic signals, such as growth factors, toxins, hormones or cytokines, can induce the extrinsic pathway of apoptosis (Figure 1.10). By using agents that activate this pathway, such as etoposide, a topoisomerase II inhibitor, or staurosporine (STS), a general kinase inhibitor, the apoptotic process and the factors involved in this process can be studied.

HeLa cells which had been treated with STS prior to fixation and DAPI staining are shown in Figure 3.7. The majority of cells display some apoptotic stage as described in Figure 3.1, including chromatin condensation (indicated by arrows in Figure 3.7), and nuclear pyknosis, which involves condensed nuclei with clumped, condensed chromatin (indicated by arrowheads in Figure 3.7).

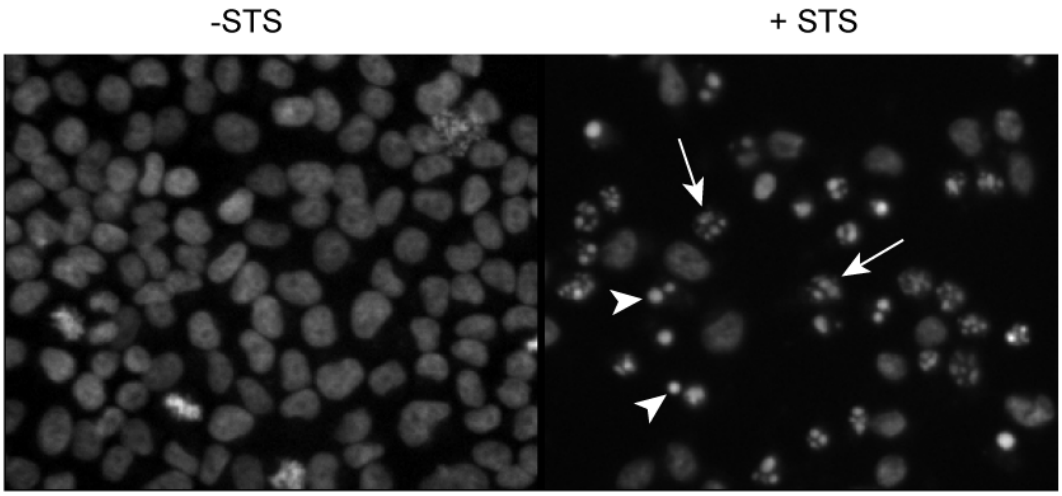


Figure 3.7. Cells treated with staurosporine (STS) undergo apoptosis. HeLa cells were treated with 1 μM STS for 4 hours before fixing, permabilising and DAPI staining. Almost 100% of cells exhibit morphology indicative of apoptosis such as cell shrinkage, chromatin condensation (indicated by arrows), membrane blebbing and nuclear pyknosis (condensed nuclei with clumped, condensed chromatin) (indicated by arrowheads).

Chromatin condensation, measured by DAPI staining, a fluorescent dye which binds strongly to DNA, advances gradually during the apoptotic process. Nuclear pyknosis, where the chromatin condensation is most visible, is one of the latter stages of apoptosis. Therefore in order to ensure cells undergoing earlier stages of apoptosis were scored correctly as apoptotic, a classification system was adopted (Figure 3.8). The stages described here are based upon previously described apoptotic stages, and

kinetic studies have shown these stages actually reflect successive steps in the apoptotic process (Daugas *et al.*, 2000). Stage I cells exhibit rippled nuclear contours and partial chromatin condensation. Stage II cells display peripheral chromatin condensation along the edge of the nuclear membrane, often leading to the formation of DAPI brightspots, while stage III cells show nuclear shrinkage and the formation of nuclear bodies (Figure 3.8). Despite this classification method it can be difficult to classify cells exactly as they may display characteristics of more than one stage, but it allows classification of cells displaying all stages of chromatin condensation, not just nuclear pyknosis.

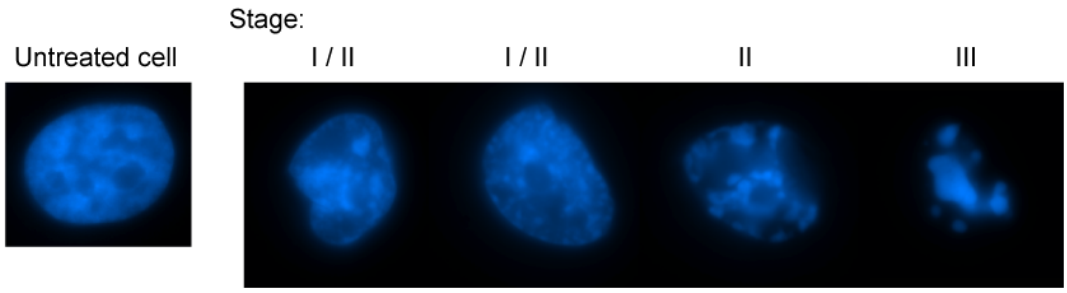


Figure 3.8. Classification of apoptotic stages. HeLa cells were treated with STS, fixed, permeabilised and DAPI stained. Images were taken of cells undergoing stages of apoptosis as defined by Daugas *et al.*, 2000. Stage I: rippled nuclear contours, partial chromatin condensation. Stage II: peripheral chromatin condensation, DAPI brightspots. Stage III: nuclear shrinkage, formation of nuclear bodies.

RNAi mediated knockdown of Acinus

In order to study the cellular effects upon Acinus knock-down, several attempts were made to achieve significant depletion of Acinus. These included using the pSUPER and pSUPERIOR vector systems (Oligoengine) and Dharmacon's siRNA reagents. Despite using Oligoengine's web-based RNAi design software to identify sequences predicted to result in substantial knockdown of Acinus, adequate depletion was not observed (data not shown). Subsequently, Dharmacon's onTarget Plus siRNA pool, containing four pre-designed RNA duplexes guaranteed to knock-down Acinus, was tested. However, the sequences of the duplexes are only available after purchase and all four duplexes were targeted against the N-terminal region of Acinus L, resulting in Acinus L knock-down, but maintained expression of Acinus S' and Acinus S (data not shown). Therefore three siRNA duplexes were custom designed against regions

common to all three isoforms of Acinus using Dharmacon's web-based siDesign centre (Figure 3.9.A). This involves use of an algorithm to design siRNAs with high functionality and specificity (Birmingham *et al.*, 2007).

A

```
                                                    siRNA 3
Acinus S    AAAGAGAGCTCGCTGCCCAAATCATTCAAGAGGAAGATCTCCGTTGTCTCAGCTACCAAG 356
Acinus S'   AAAGAGAGCTCGCTGCCCAAATCATTCAAGAGGAAGATCTCCGTTGTCTCAGCTACCAAG 331
Acinus L    AAAGAGAGCTCGCTGCCCAAATCATTCAAGAGGAAGATCTCCGTTGTCTCAGCTACCAAG 2820
            ************************************************************

                                                                siRNA 5
Acinus S    CCCCAGGTGTCAGTAGAGGTGGCCTTGCCCCCACCTGCAGAGCATGAAGTAAAGAAAGTG 776
Acinus S'   CCCCAGGTGTCAGTAGAGGTGGCCTTGCCCCCACCTGCAGAGCATGAAGTAAAGAAAGTG 751
Acinus L    CCCCAGGTGTCAGTAGAGGTGGCCTTGCCCCCACCTGCAGAGCATGAAGTAAAGAAAGTG 3240
            ************************************************************

                                        siRNA 1
Acinus S    CGAAAGGAGCAAGAAGAAGAAGAGCAAAAGGAGCGGGAGAAGGAAGCCGAGCGGGAACGG 1676
Acinus S'   CGAAAGGAGCAAGAAGAAGAAGAGCAAAAGGAGCGGGAGAAGGAAGCCGAGCGGGAACGG 1651
Acinus L    CGAAAGGAGCAAGAAGAAGAAGAGCAAAAGGAGCGGGAGAAGGAAGCCGAGCGGGAACGG 4140
            ************************************************************
```

B

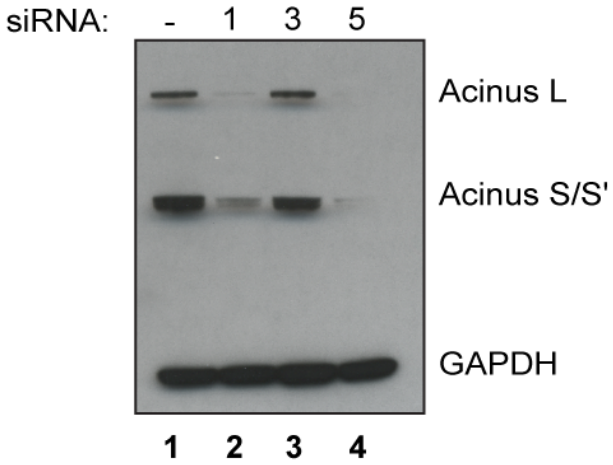


Figure 3.9. Design and testing of Acinus siRNAs. Using the web-based siDesign Center (Dharmacon) three siRNAs were designed against the region common to all three Acinus isoforms. (**A**) siRNA sequences and their alignment to Acinus isoforms. (**B**) Western blot of HeLa cell extracts which had been transfected by DharmaFECT™ 1 (Dharmacon) with the siRNAs (lanes 2-4) or mock transfected (lane 1). Extracts were blotted with an anti-Acinus antibody (Ab-2, Calbiochem) and an anti-GAPDH antibody (ab9485, AbCam). siRNA 5 was used in all future experiments and is hereafter referred to as siAcinus.

These duplexes were transfected in HeLa cells, and the relative depletion of the three isoforms of Acinus assayed by Western blot (Figure 3.9.B). The duplexes were also transfected in combination with one another (data not shown), but a greater

knockdown than that seen with siRNA5 alone was not observed. Substantial levels of Acinus depletion were observed with transfection of siRNA1 (Figure 3.9, lane 2) or siRNA5 (Figure 3.9, lane 4) compared to the mock-transfected control (Figure 3.9, lane 1), while use of GAPDH as a loading control shows similar amounts of extract were loaded. Depletion by siRNA5 was slightly better than that seen by siRNA1, therefore siRNA5 was used in all further experiments, and is hereafter referred to as siAcinus.

Over-expression of Acinus causes apoptosis

As commented previously, a proportion of cells over-expressing Acinus S displayed chromatin condensation and a loss of nuclear localisation of Acinus (Figures 3.4.C, 3.10). To investigate this effect further, and to ascertain if this is a concentration-dependent effect of Acinus, HeLa cells were transfected with low and high levels of T7-Acinus S and the number of cells exhibiting chromatin condensation counted, as assayed by DAPI staining. Cells were classified as displaying normal morphology, the apoptotic stages described above, or as undergoing mitosis (Figure 3.11.A). During prophase of mitosis the chromosomes condense in order to facilitate segregation of sister chromatids into daughter nuclei. These cells can easily be mis-classified as apoptotic, so care was taken to corectly score cells.

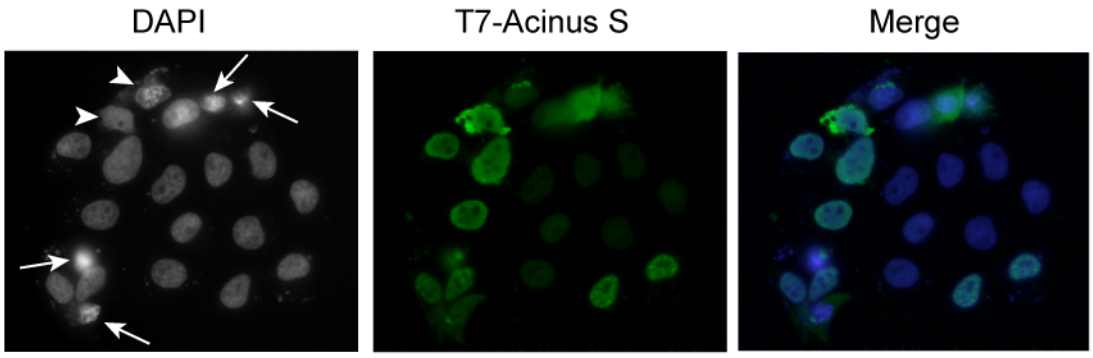


Figure 3.10. Overexpression of Acinus results in cells displaying morphological changes associated with apoptosis. HeLa cells expressing T7-Acinus S were immuno-stained with an anti-T7 antibody (Novagen) followed by an anti-mouse antibody conjugated to FITC. Approximately 35% of cells transfected with 2 µg/ml (high) and 10% of cells transfected with 20 ng/ml (low) of T7-Acinus S showed signs of apoptosis (see Figure 3.11). Arrows indicate cells with condensed chromatin, arrowheads indicate cells with rippled nuclei and peripheral chromatin condensation.

This analysis was repeated for cells depleted of Acinus by RNAi. In this case siAcinus transfected cells were also treated with STS to induce apoptosis, in order to determine if depletion of Acinus inhibited chromatin condensation during the apoptotic process.

Non-transfected and mock-transfected cells were used as controls in this assay. Cells transfected with T7-hnRNP A1 were also analysed to ensure that the effect seen was not a result of over-expression of an RNA-binding protein.

Plotting the percentage of cells exhibiting nuclear morphology characteristic of apoptosis shows that over-expression of Acinus S causes an increased number of cells to undergo apoptosis (Figure 3.11). Approximately 4% of non-transfected, mock-transfected or hnRNP A1 over-expressing cells display signs of apoptosis. This is compared to 10.5% of cells over-expressing Acinus S at low levels, and 35% of cells over-expressing Acinus S at high levels, displaying apoptosis. Statistically, this effect of over-expression of Acinus S at both low and high levels is significantly different to mock-transfected cells, when subjected to Chi-square analyses. However, almost 100% of non-transfected or mock-transfected cells undergo apoptosis after treatment with STS. Over-expression of Acinus S does not result in apoptosis on this scale.

Surprisingly, depletion of Acinus by RNAi had no effect on the number of cells displaying chromatin condensation after apoptosis induction by STS.

These results suggest that Acinus causes apoptosis in a concentration-dependent manner, which is not a general characteristic of an RNA-binding protein, but depletion of Acinus has no effect on apoptotic progression. Therefore Acinus is likely to be a contributory factor in the apoptotic pathway but is not essential, and its functions redundant.

A

	Normal	Stage I	Stage II	Stage III	Cell cycle event	
NTC	937 **94.8%**	23 **2.3%**	4 **0.4%**	3 **0.3%**	21 **2.1%**	n=988
NTC +STS	0 **0%**	36 **15.6%**	117 **50.6%**	78 **33.8%**	0 **0%**	n=231
Mock	285 **95.3%**	10 **3.3%**	3 **1.0%**	0 **0%**	1 **0.3%**	n=299
Mock +STS	3 **0.7%**	215 **51.4%**	96 **23.0%**	103 **24.6%**	1 **0.2%**	n=418
siAcinus	454 **91.3%**	29 **5.8%**	8 **1.6%**	0 **0%**	6 **1.2%**	n=497
siAcinus +STS	0 **0%**	199 **47.2%**	108 **25.6%**	115 **27.3%**	0 **0%**	n=422
T7-Acinus S (Low)	426 **89.5%**	36 **7.6%**	3 **0.6%**	11 **2.3%**	0 **0%**	n=476
T7-Acinus S (High)	314 **64.0%**	99 **20.2%**	40 **8.1%**	33 **6.7%**	5 **1.0%**	n=491
T7-hnRNP A1 (High)	192 **93.2%**	8 **3.9%**	1 **0.5%**	0 **0%**	5 **2.4%**	n=206

B

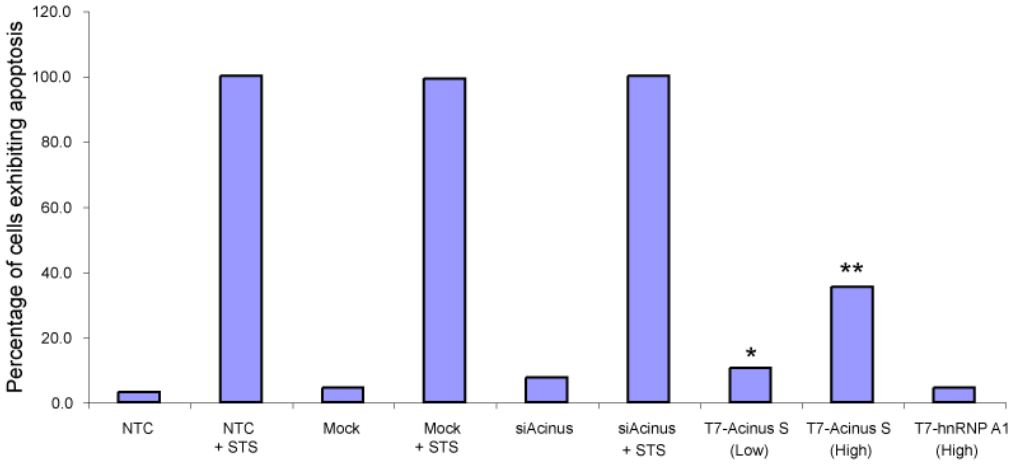


Figure 3.11. Overexpression of Acinus causes apoptosis. HeLa cells were treated as shown, fixed, permeabilised and stained with DAPI. Cells were counted and classified into the apoptotic stages previously described. (**A**) Table showing raw data counts and percentages. (**B**) Graph plotting the percentage of cells exhibiting apoptosis (cells classified into stages I, II and III were summed to give the total percentage of cells displaying apoptosis, and cell cycle events were excluded for this analysis). * Very significant: p-value of 0.002, ** highly significant: p-value of <0.0005 compared with mock transfected by Chi-square analyses. NTC: non-transfected control; STS: staurosporine.

To confirm this result by means of an alternative assay, I determined the effect of Acinus depletion or Acinus S over-expression on cell viability (Figure 3.12). This was performed using a 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfopheny)-2H-tetrazolium (MTS) assay, whereby MTS tetrazolium is converted to a formazan product by metabolically active cells and the quantity of this product can be measured by the amount of 490 nm absorbance, which is directly proportional to the number of living cells. Over-expression of Acinus S resulted in a decrease in the number of viable cells, again in a concentration-dependent manner (Figure 3.12). Statistical analysis of these results by a Student's t-test showed the differences in cell viability between mock-transfected cells, and those transfected with T7-Acinus S were significant. Depletion of Acinus by RNAi showed a slight decrease in cell viability compared to mock-transfected cells (Figure 3.12) but statistical analyses showed this was not a significant difference.

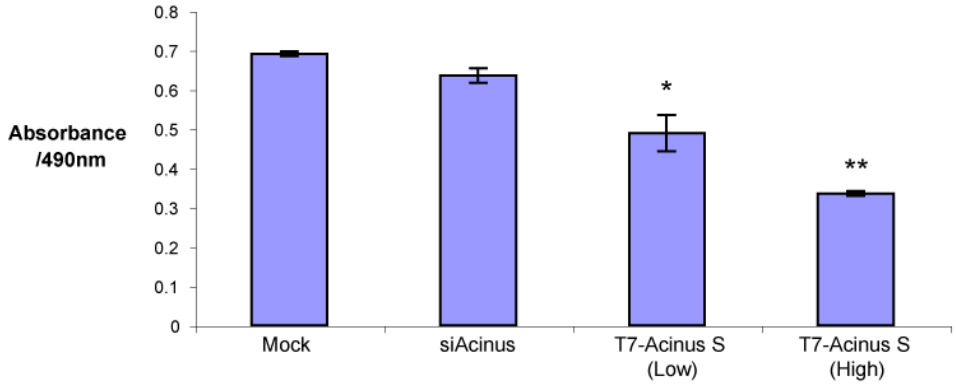


Figure 3.12. Overexpression of Acinus decreases cell viability. HeLa cells were treated as indicated and following incubation for 48 h cell viability was determined using an MTS assay (CellTiter96 Aqueous One Solution Cell Proliferation Assay, Promega). The formazan product of MTS tetrazolium converted by metabolically active cells was measured by the absorbance at 490 nm, which is directly proportional to the number of living cells. Results represent the mean ± SD, n=3 (Three biological replicates, three readings of each). * Significant: p-value of <0.02, ** highly significant: p-value of <0.0005 compared with mock transfected by Student's t-test.

The results obtained from these two independent assays demonstrates Acinus is a factor involved in mediating chromatin condensation in the apoptotic pathway, and this effect is concentration dependent. However, depletion of Acinus does not inhibit chromatin condensation during the apoptotic process or affect cell viability. Therefore Acinus is not necessary for this process to occur and its functions must

overlap with other factors acting in the same pathway, such as AIF as described above (Susin *et al.*, 1999).

3.3. Discussion

The results described above provide further evidence that Acinus is a factor involved in RNA processing. I demonstrate that the endogenous, and transiently transfected, Acinus proteins are able to bind mRNA (Figure 3.6.B). All Acinus isoforms contain a region homologous to the RRM of the Drosphila splicing regulator Sxl (Figure 3.6.A), and an RS domain. It is probable that Acinus binds RNA via the RRM, however, it is also possible that the RS domain may be responsible for the RNA-binding activity. Indeed, the RS domain of Acinus has been shown to associate directly with nuclear RNA in a non-sequence specific manner (Nikolakakai *et al.*, 2008).

T7-Acinus (987-1093), comprising the caspase-3 cleaved protein product (Figure 1.12), has been cloned into the pCGT7 vector, resulting in a T7-tagged protein. This was used in *in situ* UV crosslinking and denaturing oligo dT selection assays to determine if the RRM alone can bind mRNA. However, due to expression problems a conclusive answer as to whether this part of the protein can bind RNA was not ascertained.

Acinus was shown to display a nuclear localisation (Figure 3.2.B) and co-localisation studies with SC35, a speckle marker, showed Acinus was present in nuclear speckles (Figure 3.3). However, Acinus expression was also detected outside of these regions. Nuclear speckles are known to act as temporary storage sites of splicing factors, before these factors are recruited to active transcription sites where they can act in the pre-mRNA splicing process. The RS domain of certain SR proteins has been shown to be necessary and sufficient for targeting these factors to nuclear speckles (Li and Bingham, 1991; Hedley *et al.*, 1995; Caceres *et al.*, 1997). Therefore, the presence of an RS domain in Acinus may explain its observed localisation to nuclear speckles. Transient transfection of T7-Acinus S caused an accumulation of Acinus S in speckle-like bodies (Figure 3.4.B). Co-localisation studies were not performed with SC35, due to both available antibodies for the T7

epitope tag and SC35 being mouse monoclonal. The enlargement of the speckles may be due to increased nuclear levels of Acinus, or may be the result of inhibition of transcription, as speckles become rounded up and enlarged upon disruption of transcription or splicing (Spector *et al.*, 1983; O'Keefe *et al.*, 1994). This may reflect the role of Acinus in apoptosis which will be discussed later.

Therefore the structural, functional and sub-cellular localisation properties of Acinus all strongly imply a biological role in pre-mRNA splicing.

Many splicing factors are tissue restricted, however, Acinus was found to be ubiquitously expressed at the RNA level (Figure 3.5) and the protein can be detected in a number of cell-lines (data not shown). This may suggest Acinus is involved in constitutive splicing, or regulates alternative splicing upon induction of apoptosis, during processes such as neurodegeneration where the time taken to invoke the apoptotic pathway is on such a scale that can be regulated by alternative splicing. Identification of the transcripts to which Acinus binds, and any nearby alternative splicing events it may regulate, will elucidate the role of Acinus as a splicing factor further. To this end I used the novel cross-linking and immunopreciptitaion (CLIP) technique (Chapter 4).

The results provided in this chapter also demonstrate that Acinus is involved in the apoptotic pathway, but mechanistic details of this function could not be identified. However, I have clearly shown that over-expression of Acinus S results in apoptosis in a concentration-dependent manner (Figure 3.11), and similarly, over-expression of Acinus S decreases cell viability in a concentration-dependent manner (Figure 3.12). In contrast, depletion of Acinus by RNAi showed no significant effect in either assay. This conflicts with previous reports which showed depletion of Acinus prevents chromatin condensation (Hu *et al.*, 2005), although another study reported that knockdown of Acinus did not affect chromatin condensation, but rather inhibited DNA oligonucleosomal fragmentation, another nuclear hallmark of apoptosis (Joselin *et al.*, 2006). This suggests that Acinus' functions in the cell death pathway are redundant, and other factors such as CAD or AIF also mediate chromatin condensation.

It is interesting that over-expression of Acinus S causes a significant number of cells to undergo apoptosis, and this is concentration dependent. Only upon cleavage by caspase-3 has Acinus been reported to be active in the apoptotic pathway (Sahara *et al.*, 1999). This effect was assayed by two independent methods, by counting cells displaying morphology indicative of apoptosis (Figure 3.11) and determining cell viability (Figure 3.12). It should be noted that the MTS assay used to determine cell viability actually measures metabolic output as a readout of the number of cells present. Therefore it is possible that over-expression of Acinus S may disrupt cell metabolism, especially as mitochondrial function is closely linked in apoptotic progression, rather than directly affect cell viability. Despite this, over-expression of Acinus S does induce apoptosis in a significant proportion of cells (Figure 3.11.B). This is either a direct function of Acinus, or a result of increased amounts of protein increasing the probability of cleavage occurring by residual caspases. Nevertheless, it demonstrates the fine balance between life and death, and the co-ordinated regulation required to control this process.

When T7-Acinus (987-1093) is transfected into cells, T7 expression is not easily seen by immunostaining, but can be detected by Western blot (data not shown). However, a large amount of cellular debris is seen in these cell preparations which is not observed when cells are transfected with other constructs. The final stage of the apoptotic process is the formation of apoptotic bodies containing various parts of the cell. These bodies are engulfed by other cells and digested via the lysosomal pathway (Figure 3.1). It could be possible that over-expression of the cleaved Acinus fragment induces apoptosis rapidly, so once cells have been incubated, fixed and stained, those cells over-expressing Acinus (987-1093) have already undergone apoptosis and the only visible remains are apoptotic bodies.

The function of Acinus in chromatin condensation may be facilitated by being able to bind chromatin. Preliminary analyses using sucrose step-gradient purification of chromatin showed Acinus was a chromatin-associated protein (data not shown). Acinus has also been shown to be one of the non-histone protein targets methylated by the histone methyltransferase G9a (Rathert *et al.*, 2008). Furthermore, methylated Acinus is bound by HP1, a protein strongly associated with chromatin and involved

in heterochromatin formation. Methylation of Acinus may also allow for regulation of its activity and interaction with chromatin. Acinus also binds PKCδ and enhances its kinase activity (Hu *et al.*, 2007). If Acinus is chromatin associated, then by binding PKCδ it aids the subsequent phosphorylation of Histone H2B by placing PKCδ in close proximity to its substrate. However, although H2B phosphorylation has been linked with apoptotic cells, the mechansim by which this occurs is unknown.

Therefore, performing CLIP with Acinus serves two purposes; to identify transcripts Acinus may regulate the splicing of, and to determine if any transcripts may elucidate the role of Acinus in apoptosis.

Chapter 4: Acinus CLIP

4.1. Introduction

The previous chapter provided further evidence that Acinus has a role in RNA processing and may function in pre-mRNA splicing. Here, I use CLIP to identify *in vivo* RNA targets of Acinus.

RNA binding proteins play a central role in the post-transcriptional regulation of gene expression but often little is known about the endogenous transcripts to which they bind. Identification of these targets can provide an overview of the biological function of the protein. CLIP has the potential to identify a catalogue of *in vivo* mRNA targets to assist in understanding how these proteins regulate gene expression.

Discovering the targets to which Acinus binds may elucidate its role in both RNA processing and apoptosis, and could possibly link the two if some of its targets are transcripts encoding key apoptotic regulators. CLIP allows the identification of all RNA targets, not just protein-coding mRNAs. Therefore Acinus may bind non-coding RNAs (ncRNAs) which could bind chromatin and facilitate chromatin condensation, in a similar manner to that seen by the long ncRNA Xist which results in transcriptional repression of the inactive X chromosome by maintaining the chromatin in a heterochromatic formation, thereby reducing accessibility by the transcription machinery (Pauler *et al.*, 2007).

4.2. Cross-linking and immuno-precipitation (CLIP) technique

CLIP was developed in order to identify *in vivo* RNA targets of a neuronal splicing factor, Nova (Ule *et al.*, 2003; Ule *et al.*, 2005), and has allowed the generation of an RNA map to predict splicing regulation dependent on this protein (Ule *et al.*, 2006).

Briefly, CLIP involves an *in vivo* photo cross-linking step to capture the protein-RNA interactions, followed by partial RNase digestion to generate RNA tags of approximately 60 nt and specific immuno-precipitation for the protein of interest. By radiolabelling the RNA, the protein-RNA complexes can be separated by gel

electrophoresis whereby the complexes migrate at a higher molecular weight than that of non-cross-linked protein. The RNA can be isolated, reverse transcribed and sequenced by making use of RNA linker sequences (Figure 4.1).

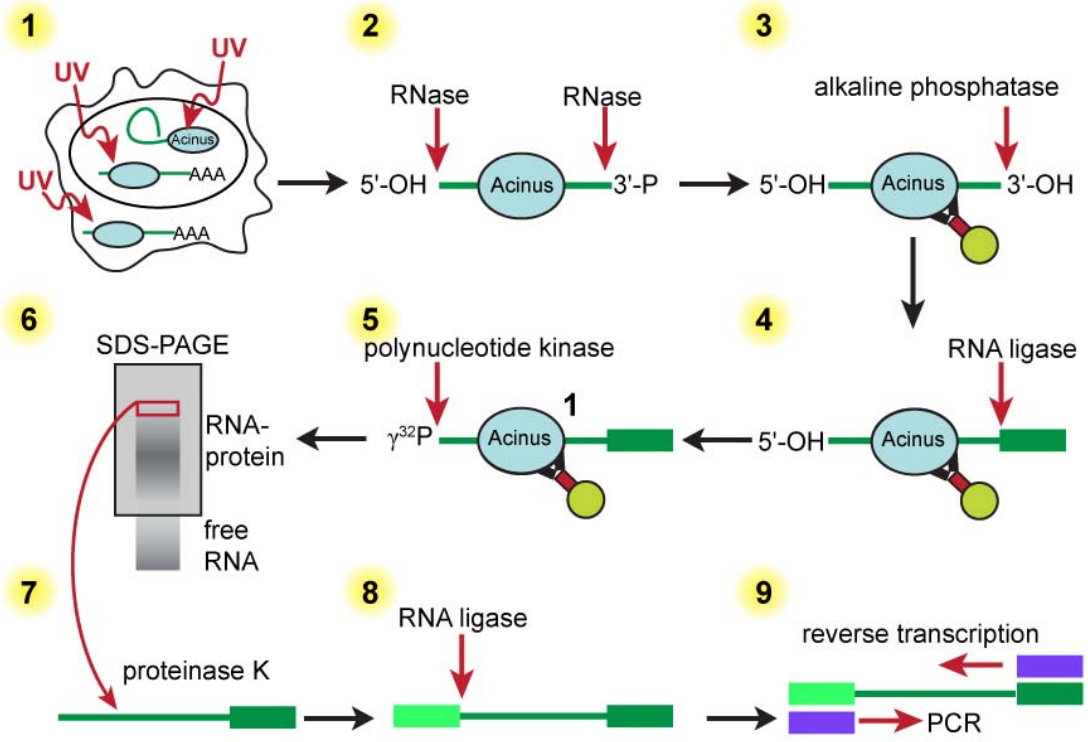


Figure 4.1. CLIP (Cross-Linking and Immuno-Precipitation) assay. Schematic of the CLIP assay for identification of protein binding sites on RNAs in the context of intact cells (adapted from Ule *et al.*, 2005). Cells are UV-irradiated to induce protein-nucleic acid cross-links (**1**). After cellular lysis the RNA is partially digested (**2**) and the protein-RNA complexes purified by immuno-precipitation (**3**). RNA is dephosphorylated (**3**), ligated to 3' RNA linker (**4**) and radioactively labelled by 5' γ^{32}P (**5**) before SDS-PAGE electrophoresis and transfer to nitrocellulose membrane (**6**). The membrane is exposed to X-ray film and the region of the membrane corresponding to protein-RNA complexes of the correct size is excised, digested by proteinase K (**7**) and 5' RNA linker ligated to the free RNA (**8**). The CLIP tags can then be amplified by RT-PCR using primers complementary to the RNA linkers (**9**).

By using an *in vivo* photo cross-linking step, protein-nucleic acid interactions are preserved in an intact cell. Formaldehyde cross-linking can also be used to capture protein-nucleic acid interactions. However, this can induce large multi-molecular chemical bridges which can create artefacts, and make it difficult to identify direct versus indirect protein-nucleic acid interactions. Conversely, UV cross-linking takes advantage of the natural photoreactivity of nucleic acid bases, especially

pyrimidines, and specific amino acids such as cysteine, lysine, tryptophan and tyrosine. UV-induced cross-links only occur between proteins and nucleic acids, protein-protein cross-links cannot occur. Therefore using this approach one can be certain that the protein-nucleic acid interaction is specific and the site can be mapped accurately. The covalent bond formed by UV cross-linking is irreversible and therefore further manipulations can be performed, such as the RNA degradation step without concern for loss of the relevant portion of the RNA which remains bound to the protein of interest.

Previously, techniques such as SELEX and RIPs, described in Chapter 1, have been used to determine binding sequences or specific substrates of RNA binding proteins. CLIP offers two main advantages over these techniques; the ability to identify *in vivo* targets, and the use of UV-crosslinking prevents artefacts commonly associated with formaldehyde crosslinking.

4.3. Results

Optimisation of CLIP for Acinus

It is necessary to optimise CLIP for each protein studied. CLIP was already being used in the lab for SF2/ASF and hnRNP A1, however I changed a number of steps in order to obtain acceptable results for Acinus. These are detailed in Table 4.1 and involve steps prior to, and including the immuno-precipitation reaction, such as lowering the dose of UV-irradiation to prevent protein degradation, using less reactive RNases to obtain a CLIP tag of an appropriate size, and increasing the denaturing conditions and stringency of the immuno-precipitation to ensure only specific protein-RNA complexes were purified.

CLIP was performed with endogenous Acinus proteins, as well as transiently transfected T7-Acinus S and T7-Acinus (987-1093) (Figure 1.12). However, the immuno-precipitation reactions proved cleaner with the T7-tagged proteins, and due to complexities of T7-Acinus (987-1093) inducing wide-spread apoptosis, the results presented here are targets found to bind to T7-Acinus S.

Step	Change	Reason/Benefit
UV-irradiation (1)	Titrating the dose of UV-irradiation over a range of 200-400 mJ/cm2 showed protein-RNA cross-linking still occurred at 300 mJ/cm2. This lower dose (compared to the previously used dose of 400 mJ/cm2) was used.	UV-irradiation induces apoptosis and causes protein degradation. Using a lesser dose helps to preserve intact proteins and is especially important when studying a protein, such as Acinus, which is cleaved in the apoptotic process.
RNA digestion (2)	The RNase A/T1 cocktail was replaced with RNase T1 alone.	RNase A hydrolyses RNA at C and U residues, while RNase T1 hydrolyses RNA at G residues. The Acinus CLIP tags produced when partially digesting with RNase A/T1 were consistently short; therefore the less reactive RNase T1 was used to generate longer CLIP tags of a more informative size.
Immuno-precipitation (3)	The concentration of SDS in the lysis buffer was raised to 0.2%.	With lower concentrations of SDS a band was observed on the autoradiogram in the absence of UV-irradiation. Protein-RNA complexes are unlikely to be seen in the absence of cross-linking, therefore this was possibly due to a kinase associating with Acinus and phosphorylating it with the radio-labelled phosphate. Increasing the SDS concentration, and therefore preventing the formation of protein-protein interactions eliminated this band.
Immuno-precipitation (3)	Following the immuno-precipitation the beads were washed in lysis buffer containing 400 mM salt.	This high salt wash increased the stringency of the immuno-precipitation and prevented the protein from sticking to the beads alone.

Table 4.1. Optimisation of Acinus CLIP. The details and reasons of the major changes to the CLIP protocol used for Acinus as compared to that used for hnRNP A1 or detailed in Ule *et al.*, 2005. The numbers refer to those depicted in Figure 4.1.

Figure 4.2 shows UV-dependent, RNase-sensitive complexes immunoprecipitated by the anti-T7 monoclonal antibody. A range of RNase concentrations are required to determine if the protein is bound to RNA. Over-digestion results in a band close to the molecular weight of the RNA-binding protein, while under-digestion causes an upward shift in the molecular weight of the complex. The average molecular weight of 50 nt of RNA is ~16 kDa, with the addition of the linker regions, the region to be excised is about 20 kDa above the expected molecular weight of the protein. Figure 4.2 shows the analysis of CLIP protein-RNA complexes for both T7-Acinus (987-1093) (A and B) and T7-Acinus S (C and D). The RNA was extracted from the membrane and amplified by RT-PCR. BLAST was used to align sequences to the human genome allowing for two mis-matches or gaps within the alignment.

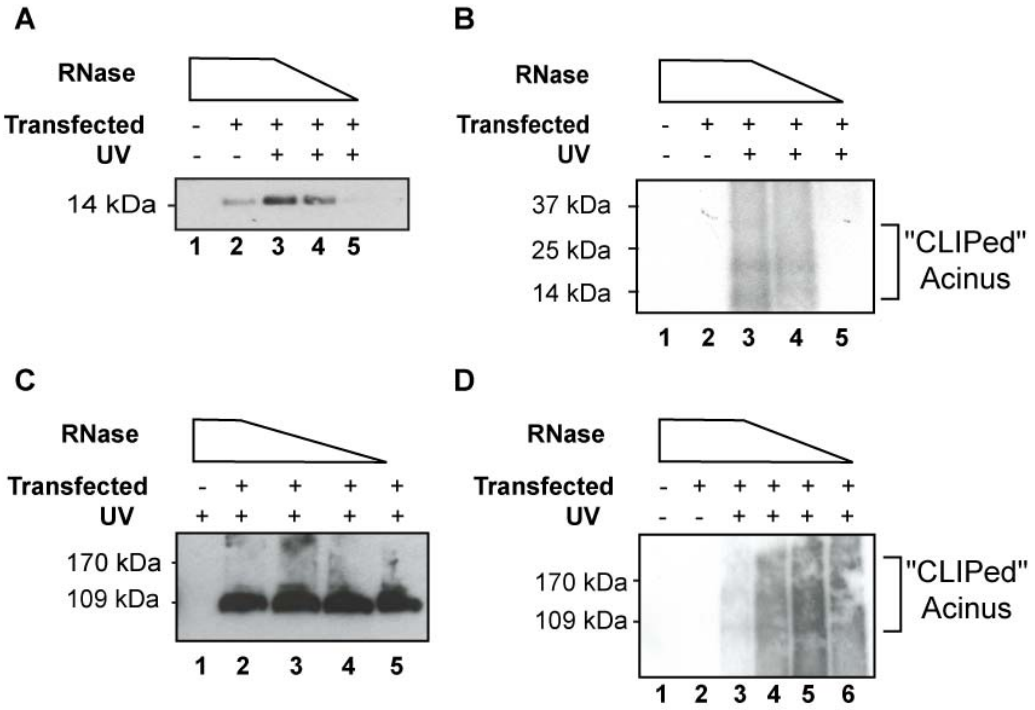


Figure 4.2. Analysis of Acinus-RNA complexes by CLIP. Complexes were isolated with anti-T7 tag immunoprecipitation from 293T cells transfected with T7-Acinus (987-1093) (**A**, **B**) or T7-Acinus S (**C**, **D**). (**A**, **C**) Western blot analysis of the immunoprecipitation reaction to show the mobility of the proteins not bound to RNA. (**B**, **D**) Analysis of SDS-PAGE separated protein-RNA complexes after RNA digestion and T4 PNK assay.

Identification of Acinus targets

To determine if the endogenous binding sites of Acinus can elucidate the role of Acinus in alternative splicing, the transcripts targeted by Acinus were manually curated for alternative splicing events using Ensembl, UCSC genome browser and Fast-DB databases.

Of the 43 binding sites present in protein-coding genes, 13 (30%) of these were located in, or adjacent to, an alternative splicing event (Tables 4.2-4.7, Figure 4.3). These include the majority of possible alternative events (Figure 1.5), with Acinus binding sites located in an alternative cassette exon (NBPF14), in exons with alternative 3' splice sites (MZF1 and SFRS15), in exons with alternative 5' splice sites (FNBP4 and SFRS10), and in exons with retained introns (GALNS and MYST1) or internal exon deletions (GANAB and PEX6). Acinus binding sites were also found adjacent to alternative cassette exons (CACHD1, MZF1, UBE2I and RCL1). The presence of an Acinus binding site adjacent to alternative splicing events could suggest a role for Acinus in the regulation of alternative splicing of these transcripts.

Several of the alternative splicing events lead to PTCs and therefore are potential substrates for NMD (MZF1, SFRS15 and GALNS), but some events alter the primary structure of the encoded polypeptide (NBPF14, SFRS10 and MYST1) (Table 4.8).

The remainder of the protein-coding transcripts bound by Acinus are listed in Table 4.9. See Appendix 1 for full sequence details of the CLIP tag, genomic co-ordinates and Ensembl gene ID.

The protein product of a number of the transcripts bound by Acinus have been reported to function in RNA processing (SFRS10 and SFRS15), chromatin binding (MYST1 and SIRT3) or apoptosis (ACIN1, MZF1 and F2R).

Gene ID	Exon Position	Gene Description	Function
NBPF14	12	Neuroblastoma breakpoint family member 14	Unknown

Table 4.2. Alternative cassette exons bound by Acinus. Exons were annotated manually according to the FastDB database (www.fast-db.com, de la Grange *et al.*, 2005). Column headers: Gene ID, official gene symbol; Exon/Intron Position, number of exon/intron within the primary transcript; Gene Description; Function, annotated function of the protein product.

Gene ID	Exon Position	Gene Description	Function
MZF1	9	Myeloid zinc finger 1	Transcription regulation
SFRS15	8	Splicing factor, arginine/serine-rich 15	mRNA processing

Table 4.3. Acinus target exons containing alternative 3' splice sites.

Gene ID	Exon Position	Gene Description	Function
FNBP4	1	Formin-binding protein 4	Cell adhesion; G-protein coupled receptor activity
SFRS10	4	Splicing factor, arginine/serine-rich 10 (transformer 2 homolog, Drosophila)	mRNA processing

Table 4.4. Acinus target exons containing alternative 5' splice sites.

Gene ID	Exon Position	Gene Description	Function
GALNS	3	Galactosamine (N-acetyl)-6-sulfate sulfatase	Hydrolase activity
GANAB	25	Glucosidase, alpha; neutral AB	Hydrolase activity
MYST1	11	MYST histone acetyltransferase 1	Chromatin assembly or disassembly; Transcription regulation
PEX6	17	Peroxisomal biogenesis factor 6	Peroxisome organization and biogenesis

Table 4.5. Acinus target exons containing a retained intron.

Gene ID	Exon Position	Gene Description	Function
CACHD1	8	Cache domain- containing protein 1	Calcium ion transport
MZF1	9	Myeloid zinc finger 1	Transcription regulation
UBE2I	1	Ubiquitin-conjugating enzyme E2I	Ubiquitin cycle; Chromosome segregation

Table 4.6. Acinus target exons located adjacent to alternative cassette exons.

Gene ID	Intron Position	Gene Description	Function
RCL1	1	RNA terminal phosphate cyclase-like 1	RNA-3'-phosphate cyclase activity

Table 4.7. Acinus target introns located adjacent to alternative cassette exons.

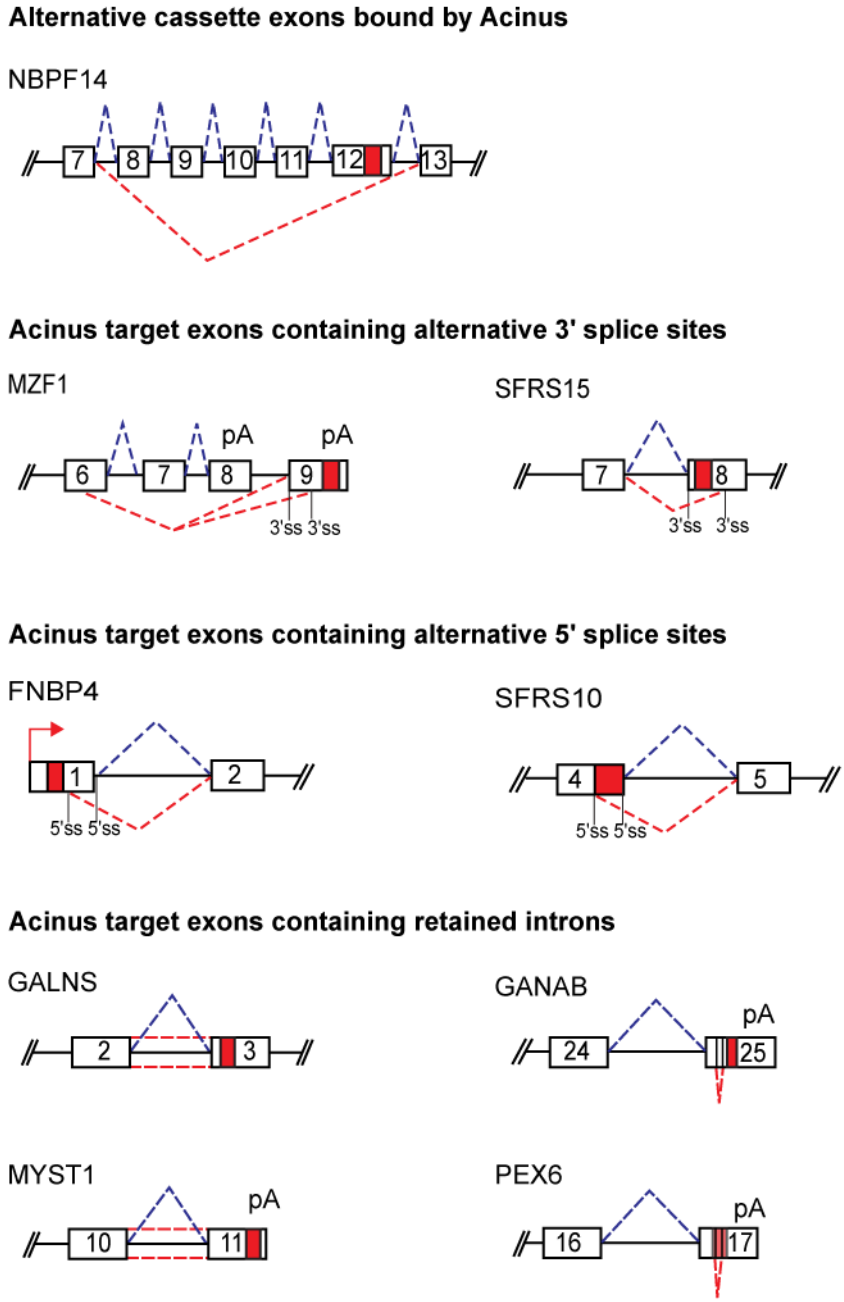


Figure 4.3. RNA targets of Acinus identified by CLIP which are located in or nearby sites of alternative splicing events. Red box indicates CLIP tag location within the transcript. Splicing profiles determined from FastDB (www.fast-db.com, de la Grange *et al.*, 2005) or UCSC (http://genome.ucsc.edu, Kent *et al.*, 2002) databases. pA: polyadenylation site; ss: splice site; red arrow: transcription start site.

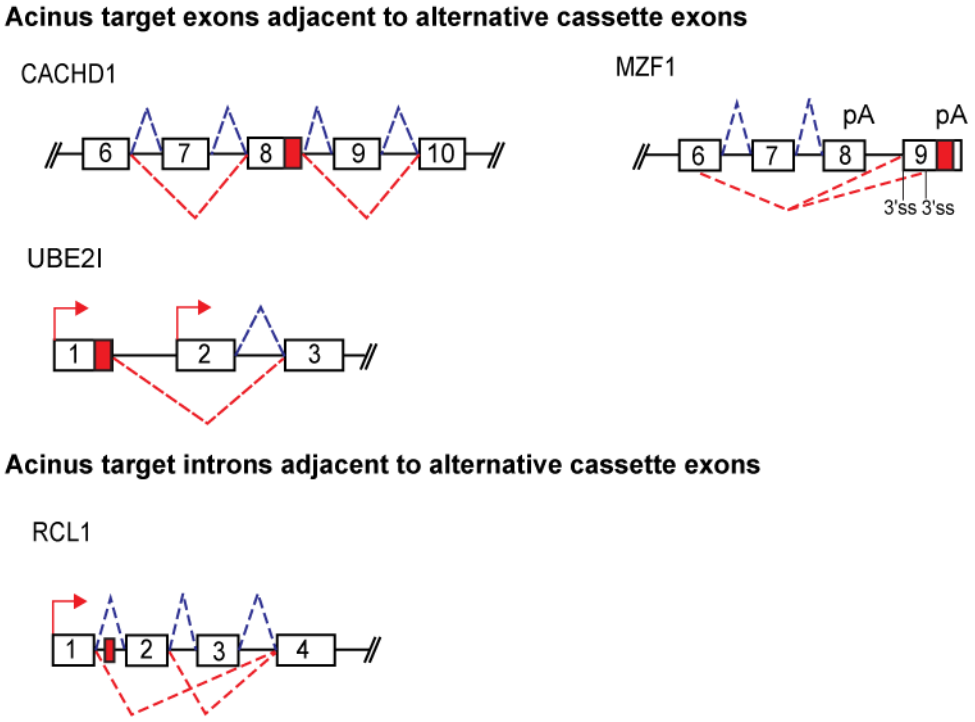


Figure 4.3. RNA targets of Acinus identified by CLIP which are located in or nearby sites of alternative splicing events (continued).

Transcript	Alternative splicing event	Effect on protein product
NBPF14	Exons 8-12 skipping	In-frame skipping event of 310 amino acids which contains repeats of unknown function.
MZF1	Alternative 3' ss in exon 9	Use of distal 3' ss introduces a PTC - possible substrate for NMD.
MZF1	Alternative 5' ss in exon 6	Use of distal 5' ss introduces a PTC - possible substrate for NMD.
SFRS15	Alternative 3' ss in exon 8	Use of distal 3' ss introduces a PTC - possible substrate for NMD.
FNBP4	Alternative 5' ss in exon 1	Use of distal 5' ss results in an in-frame insertion of 2 amino acids
SFRS10	Alternative 5' ss in exon 4	Use of proximal 5' ss results in an in-frame deletion of 21 amino acids. This occurs in part of the RS domain, but more than 10 RS dipeptides remain.
GALNS	Intron 2 retention	Introduces a PTC - possible substrate for NMD.
GANAB	Internal exon deletion in exon 25	Event occurs in 3' UTR.
MYST1	Intron 10 retention	Results in the production of two isoforms with varying C-termini.
PEX6	Internal exon deletion in exon 17	Event occurs in 3' UTR.
CACHD1	Exons 7 and 9 skipping	Event occurs in 5' UTR.
UBE2I	Exon 2 skipping	Event occurs in 5' UTR.
RCL1	Exons 2-3 skipping	Event occurs in 5' UTR.

Table 4.8. Effect on protein product of alternative splicing of Acinus bound transcripts. The effect of the described alternative splicing event was determined by manually manipulating sequences and by searching databases for the potential protein product isoforms.

Gene ID	Exon Position	Gene Description	Function
ACIN1	13,15,18, 21	Apoptotic chromatin condensation inducer in the nucleus	Apoptosis
ANAPC1	7,8	Anaphase-promoting complex subunit 1	Cell cycle regulation
ATP5B	7	ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide	ATP binding; Transporter activity
CACNA1D	1	Calcium channel, voltage-dependent, L type, alpha 1D subunit	Ion transport; G-protein signalling
CCDC66	5	Coiled-coil domain containing 66	Unknown
CDK6	2	Cyclin-dependent kinase 6	Protein-kinase activity
DIS3	23	DIS3 mitotic control homolog	Exonuclease activity
DLX4	1	Distal-less homeobox 4	Transcription regulation
DNAH11	82	Dynein, axonemal, heavy chain 11	ATPase activity; Microtubule-based movement
EEF1B2	5	Eukaryotic translation elongation factor 1 beta 2	Translational elongation
FEM1A	2	Fem-1 homolog a (C. elegans)	Receptor activity
F2R	2	Coagulation factor II (thrombin) receptor	Cell cycle regulation; Apoptosis
GALNT7	3	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 7 (GalNAc-T7)	Protein glycosylation
GNL3	6	Guanine nucleotide-binding protein-like 3	Cell cycle regulation
LTBP4	15,16	Latent transforming growth factor beta binding protein 4	Transmembrane receptor activity
PDE1A	11	Phosphodiesterase 1A, calmodulin-dependent	Signal transduction
PQBP1	8	Polyglutamine binding protein 1	Transcription regulation
PSMB1	3	Proteasome (prosome, macropain) subunit, beta type, 1	Peptidase activity
RBPJ	10	Recombination signal binding protein for immunoglobulin kappa J region	DNA recombination; Transcription regulation
SEC13	9	SEC13 homolog (S. cerevisiae)	Intracellular protein transport
SRD5A1	5	Steroid-5-alpha-reductase, alpha polypeptide 1 (3-oxo-5 alpha-steroid delta 4-dehydrogenase alpha 1)	Lipid metabolism
SIRT3	8	Sirtuin (silent mating type information regulation 2 homolog) 3 (S. cerevisiae)	Chromatin silencing; Transcription regulation
SSBP2	16	Single-stranded DNA binding protein 2	Transcription regulation
TMPO	1	Thymopoietin	Glycolysis; Transcription regulation
TNPO3	20	Transportin 3	Protein transport
WDR60	13,14	WD repeat domain 60	Unknown
ZFYVE20	9	Zinc finger, FYVE domain containing 20	Protein transport
BPHL	Intron 7	Biphenyl hydrolase-like	Proteolysis

Table 4.9. Other Acinus target transcripts.

Analysis of sites to which Acinus binds

The motif finding algorithm MEME was used to search for over-represented sequences, however, no identifiable binding site for Acinus could be determined.

Acinus is a member of the EJC (Figure 1.4), a complex which is deposited at a specific location, about 20-24 nt upstream of the exon-exon junction on every pre-mRNA (Le Hir *et al.*, 2000b; Le Hir *et al.*, 2003). Therefore if Acinus is a binding component of the EJC, it may bind RNA in a non-specific fashion.

The positions of the sequences bound by Acinus were analysed with respect to exon-exon junctions, to determine if they overlapped the region of EJC deposition (Table 4.10). Only 11 out of 43 (26%) sequences analysed overlapped the EJC site. Therefore it is unlikely that Acinus acts as a core binding protein of the EJC, but is rather a nuclear-associated outer core component as previously described (Tange *et al.*, 2005).

CLIP genomic coordinates	Gene ID	Exon position	Exon coordinates	Exon ID	Distance from 3' end of exon /nt	EJC site
chr14:22,603,447-22,603,500	ACIN1	13	chr14:22,603,404-22,603,537	ENSE00000653848	90-37	N
chr14:22,602,590-22,602,624	ACIN1	15	chr14:22,602,485-22,602,626	ENSE00000653845	36-2	Y
chr14:22,601,253-22,601,319	ACIN1	18	chr14:22,601,217-22,601,324	ENSE00000653842	71-5	Y
chr14:22,598,412-22,598,445	ACIN1	21	chr14:22,597,616-22,598,523	ENSE00001031613	111-78	N
chr2:112,342,135-112,342,144	ANAPC1	7	chr2:112,342,071-112,342,144	ENSE00001146353	9-0	N
chr2:112,338,854-112,338,866	ANAPC1	8	chr2:112,338,854-112,338,999	ENSE00001146346	145-133	N
chr12:55,322,589-55,322,620	ATP5B	7	chr12:55,322,509-55,322,631	ENSE00001181239	42-11	Y
chr1:64,867,710-64,867,752	CACHD1	5	chr1:64,867,626-64,867,752	ENSE00001361287	42-0	Y
chr3 53,504,274-53,504,295	CACNAID	1	chr3:53,504,116-53,504,300	ENSE00001365080	26 5	Y
chr3:56,572,949-56,572,968	CCDC66	5	chr3:56,572,752-56,573,193	ENSE00001436608	244-225	N
chr7:92,300,701-92,300,720	CDK6	2	chr7:92,300,341-92,300,940	ENSE00001132651	239-220	N
chr13:72,229,735-72,229,753	DIS3	21	chr13:72,227,543-72,232,017	ENSE00001355717	2282-2264	NA
chr17:45,401,725-45,401,746	DLX4	1	chr17:45,401,561-45,402,114	ENSE00001293328	389-368	N
chr7:21,907,645-21,907,695	DNAH11	83	chr7:21,907,150-21,907,982	ENSE00001416067	337-287	NA
chr2:206,735,019-206,735,052	EEF1B2	5	chr2:206,735,007-206,735,073	ENSE00000784966	54-21	Y
chr5:76,064,615-76,064,635	F2R	2	chr5:76,063,895-76,067,054	ENSE00001259490	2439-2419	NA
chr19:4,743,870-4,743,893	FEM1A	1	chr19:4,742,728-4,746,571	ENSE00000951494	2701-2678	N
chr11:47,745,377-47,745,398	FNBP4	1	chr11:47,745,197-47,745,569	ENSE00001506181	192-171	N
chr16: 87,435,825-87,435,855	GALNS	3	chr16:87,435,806-87,435,880	ENSE00000946131	55-25	Y
chr4:174,449,868-174,449,894	GALNT7	3	chr4:174,449,834-174,450,000	ENSE00001081447	132-106	N
chr11:62,149,412-62,149,442	GANAB	25	chr11:62,148,876-62,149,972	ENSE00001418010	560-530	NA
chr3:52,698,204-52,698,246	GNL3	6	chr3:52,698,130-52,698,262	ENSE00001079830	58-16	Y
chr19:45,807,395-45,807,407	LTBP4	15	chr19:45,807,276-45,807,407	ENSE00001138350	12-0	N
chr19:45,807,492-45,807,508	LTBP4	16	chr19: 45,807,492-45,807,617	ENSE00001138344	125-109	N
chr16:31,050,059-31,050,100	MYST1	11	chr16:31,050,017-31,050,206	ENSE00001172748	147-106	N
chr19: 63,765,382-63,765,404	MZF1	6	chr19:63,765,097-63,766,683	ENSE00000648318	1301-1279	NA
chr1:146,478,379-146,478,400	NBPF14	13	chr1:146,478,304-146,478,412	ENSE00001561510	33-12	Y
chr2:182,774,700-182,774,718	PDE1A	9	chr2:182,774,660-182,774,761	ENSE00000964530	61-43	N

Table 4.10. Analysis of protein-coding Acinus CLIP tag locations in relation to exon-exon junctions.

CLIP genomic coordinates	Gene ID	Exon position	Exon coordinates	Exon ID	Distance from 3' end of exon /nt	EJC site
chr6:43,039,936-43,040,002	PEX6	17	chr6:43,039,725-43,040,187	ENSE00001359565	251-185	N
chr6:170,697,141-170,697,181	PSMB1	3	chr6:170,697,116-170,697,197	ENSE00000766144	56-16	Y
chrX:48,645,228-48,645,316	PQBP1	8	chrX:48,645,149-48,645,357	ENSE00001151029	129-41	NA
chr4:26,040,637-26,040,678	RBPJ	10	chr4:26,040,618-26,040,773	ENSE00001122452	136-95	N
chr3:10,321,729-10,321,801	SEC13	9	chr3:10,321,717-10,321,840	ENSE00001029806	111-39	N
chr3:187,127,081-187,127,109	SFRS10	2	chr3:187,127,083-187,127,216	ENSE00001029249	135-107	N
chr21:31,990,834-31,990,853	SFRS15	8	chr21:31,990,753-31,990,934	ENSE00001025679	100-81	N
chr11:205,813-205,834	SIRT3	7	chr11:205,031-206,718	ENSE00001493185	905-884	NA
chr5:6,721,930-6,721,954	SRD5A1	5	chr5:6,721,315-6,722,673	ENSE00001206808	743-719	NA
chr5:80,752,036-80,752,070	SSBP2	17	chr5:80,751,428-80,752,108	ENSE00001418561	72-38	NA
chr12:97,433,766-97,433,786	TMPO	1	chr12:97,433,536-97,434,055	ENSE00001407932	289-269	N
chr7:128,399,831-128,399,851	TNPO3	19	chr7:128,399,716-128,399,872	ENSE00001130703	41-21	Y
chr16:1,299,711-1,299,736	UBE2I	1	chr16:1,299,639-1,299,775	ENSE00001529011	64-39	N
chr7:158,398,540-158,398,548	WDR60	13	chr7:158,398,437-158,398,548	ENSE00001019332	8-0	N
chr7:158,399,685-158,399,713	WDR60	14	chr7:158,399,685-158,399,773	ENSE00001019343	88-60	N
chr3:15,098,953-15,098,991	ZFYVE20	9	chr3:15,098,878-15,099,119	ENSE00000901299	166-128	N

Table 4.10. Analysis of protein-coding Acinus CLIP tag locations in relation to exon-exon junctions (continued).
Headers: CLIP genomic coordinates; Gene ID, official gene symbol; Exon position, number of exon within the primary transcript; Exon coordinates, genomic coordinates of exon; Exon ID, ensembl exon ID; Distance from 3' end of exon/nt, the position of the CLIP tag with respect to the downstream exon-exon junction; EJC site, if the CLIP tag overlaps the deposition site of the EJC, N: No, Y: Yes, NA: Not applicable (for terminal exons where there is no downstream exon-exon junction). The EJC has been reported to be deposited about 20 -24 nt upstream of the exon-exon junction (Le Hir et al., 2000b; Le Hir et al., 2003) and has a footprint of about 8 nt. Therefore if the CLIP tag overlapped positions 26-18 from the exon-exon junction it was said to be present at the EJC deposition site.
CLIP tags which overlap two exons appear as two entries (ANAPC1, LTBP4 and WDR60).

Acinus is an alternative splicing regulator

To determine if Acinus is an alternative splicing regulator, six of the transcripts identified by CLIP where the binding site of Acinus is in, or nearby an alternative splicing event, were investigated. The levels of Acinus were manipulated in HeLa (Figure 4.4) or 293T cells (data not shown) by RNAi depletion of Acinus (Figure 4.4.A, lanes 3 and 4) or transient expression of T7-Acinus S at high (Figure 4.4.B, lanes 5 and 6) and low levels (Figure 4.4.B, lanes 7 and 8). RNA prepared from these cells was used in RT-PCR based assays, with primers designed in regions flanking the alternative processing event, to determine if the levels of Acinus modulate the splicing pattern.

In these assays it was essential to use RNA of high quality and integrity. This was analysed by on-chip electrophoresis using the Bioanalyzer 2100 (Agilent), which gives an indication of the integrity of the RNA based on the 18S and 28S ribosomal RNA ratio. It also allows for comparison between samples by use of the RNA integrity number (RIN), which is based on the entire electrophoretic trace of the RNA sample, including the presence or absence of degradation products. Figure 4.5.A shows the electropherograms and densitometry plots for each RNA sample, while other information including the RIN is shown in Figure 4.5.B. The RNA samples were also used to amplify 1 kb fragments of both Beta-actin and GAPDH using intron-spanning primers to futher assess the quality of the RNA (Figure 4.5.C and D). Both forms of analysis showed the RNA samples were of satisfactory quality to be used as substrates in RT-PCR assays.

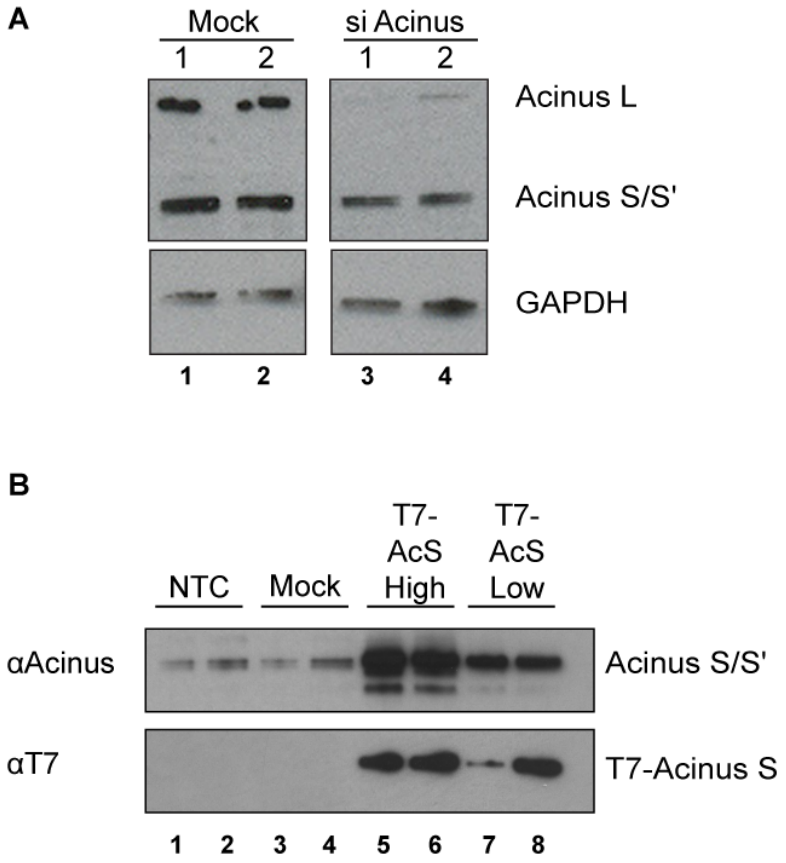
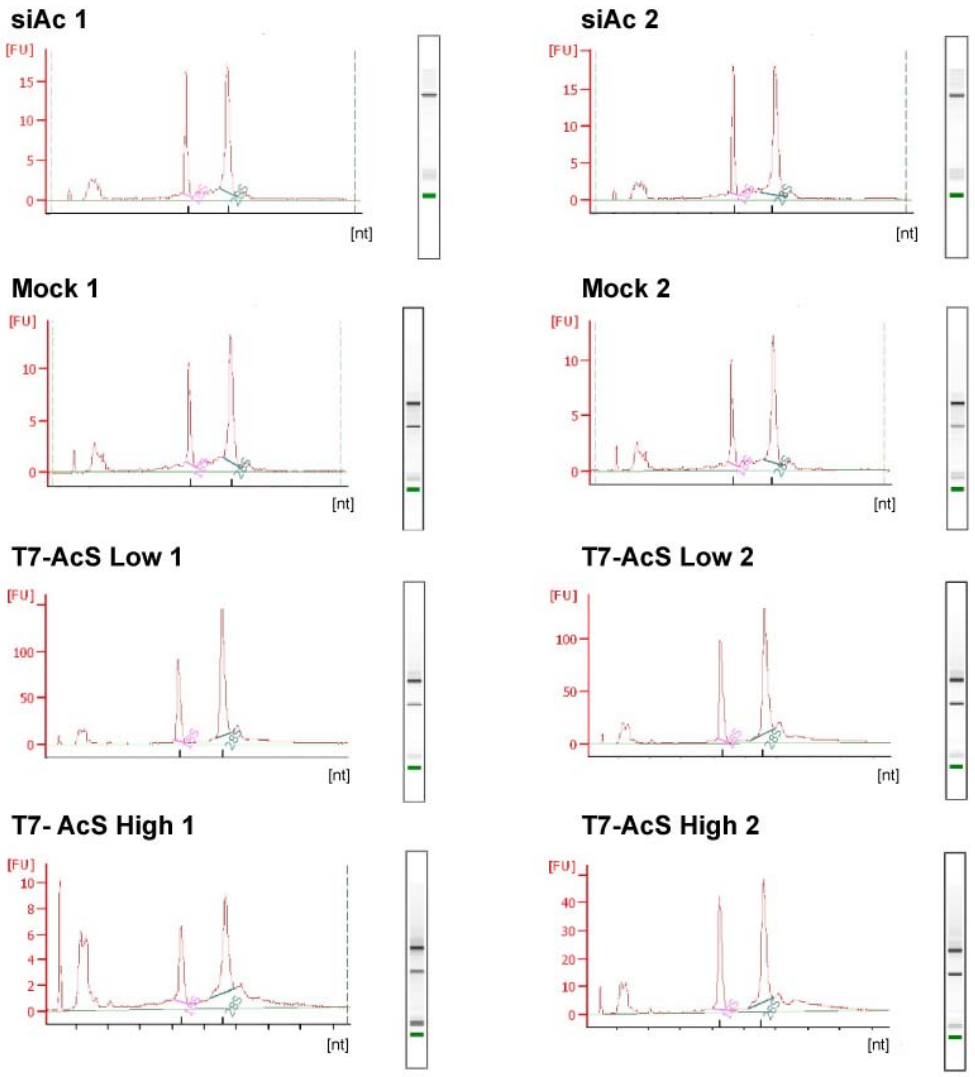


Figure 4.4. Analysis of Acinus knockdown and overexpression. HeLa cells were transfected by DharmaFECT™ 1 (Dharmacon) with siAcinus (siAc) or by lipofectamine-2000 (Invitrogen) with T7-Acinus S (T7-AcS) at two concentrations; 2 µg/ml (High) and 20 ng/ml (Low). These cells were used to prepare a protein extract to assay the depletion or over-expression of Acinus. (**A**) Western blot of HeLa cell extracts which had been mock transfected (lanes 1 and 2) or transfected with siAcinus (lanes 3 and 4). Extracts were blotted with an anti-Acinus antibody (Ab-2, Calbiochem) and GAPDH was used as a loading control (ab9485, AbCam). (**B**) Western blot of HeLa cell extracts which had been non-transfected (lanes 1 and 2), mock transfected (lanes 3 and 4), transfected with 2 µg/ml of T7-Acinus S (lanes 5 and 6) or transfected with 20 ng/ml of T7-Acinus S (lanes 7 and 8). Extracts were blotted with an anti-Acinus antibody (Ab-2, Calbiochem) (upper panel) and an anti-T7 antibody (Novagen) (lower panel). Two biological replicates are shown for each condition. NTC: Non-transfected control. (Figure 4.4.B above is identical to Figure 3.4.D, and is included here for comparison purposes).

A



B

RNA	Concentration (ng/µl)	28s:18s ratio	RNA integrity number
siAc 1	273	1.8	9.3
siAc 2	310	2.0	9.3
Mock 1	131	1.9	8.6
Mock 2	101	1.9	8.6
T7-Ac S Low 1	480	1.6	10.0
T7-Ac S Low 2	500	1.5	9.9
T7-Ac S High 1	359	1.6	8.2
T7-Ac S High 2	272	1.2	9.6

Figure 4.5. Quality analysis of RNA to be used in RT-PCR assays.

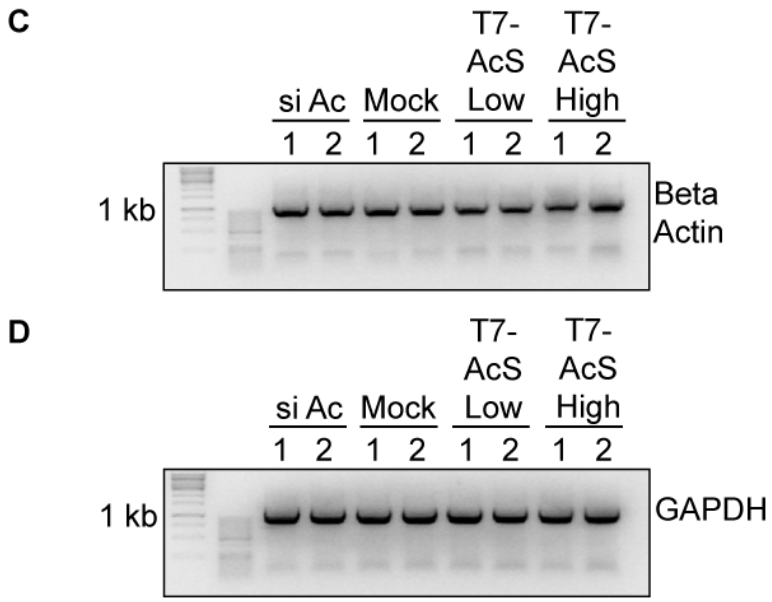


Figure 4.5. Quality analysis of RNA to be used in RT-PCR assays (continued). HeLa cells were transfected by DharmaFECT™ 1 (Dharmacon) with siAcinus (siAc) or by lipofectamine-2000 (Invitrogen) with T7-Acinus S (T7-AcS) at two concentrations; 2 µg/ml (High) and 20 ng/ml (Low). Total RNA was prepared from these cells using Trizol (Invitrogen) and analysed by on-chip gel electrophoresis using the Bioanalyzer 2100 (Agilent). The electropherograms, with the 18S and 28S rRNA peaks labelled, and densitometry plots (gel-like images) for each RNA sample are shown in **A**. The concentration, ribosomal RNA ratio and RNA integrity number (RIN) are shown in **B**. The RIN gives an estimate of the integrity of total RNA samples based on the entire electrophoretic trace of the RNA sample and allows RNA samples to be directly compared. 10 indicates a perfect RNA sample without any degradation products, whereas 1 marks a completely degraded sample. To further assess the quality of the RNA, a 1kb fragment of beta-actin (**C**) or GAPDH (**D**) were amplified by RT-PCR (Superscript™ III one-step RT-PCR sytem with Platinum® Taq, Invitrogen) from 500ng of the prepared total RNA samples (Ta 60°C, 40 cycles). Two biological replicates are shown for each condition.

Seven alternative splicing events were investigated (Figure 4.6), and of these at least two appear to be regulated by Acinus. Over-expression of Acinus S results in intron 10 retention of MYST1 (Figure 4.6.B). This modulation occurs in a concentration-dependent manner as with higher levels of Acinus S more intron retention is detected. A band corresponding to about 500 bp (marked with an asterix) was also detected upon over-expression of Acinus S. All three amplified bands were excised, cloned into pGem T-easy and 20 transformants sequenced. As predicted the sequence of the 300 nt band was that of exon 10 followed by exon 11, and the sequence of the 600 nt band matched exon 10, intron 10 and exon 11. However, screening of 20 transformants could not reveal the composition of the 500 nt band.

The sequences matched back to MYST1 but none were 500 nt in length. Testing the primers by *in silico* PCR analysis also results in no predicted product of 500 nt. Therefore although this product also seems to be under regulation by Acinus, its composition is unknown.

Over-expression of Acinus S promotes exon 7 inclusion in the CACHD1 transcript (Figure 4.6.F). It is possible over-expression of Acinus promotes exon 9 skipping of the CACHD1 transcript (Figure 4.6.G) but this result is less convincing. Acinus does not display any detectable modulation of the depicted alternative splicing events in transcripts NBPF14 (Figure 4.6.A), GANAB (Figure 4.6.C), GALNS (Figure 4.6.D) or RCL1 (Figure 4.6.E).

These results clearly show Acinus is a regulator of alternative splicing, although the precise mechanism requires further investigation.

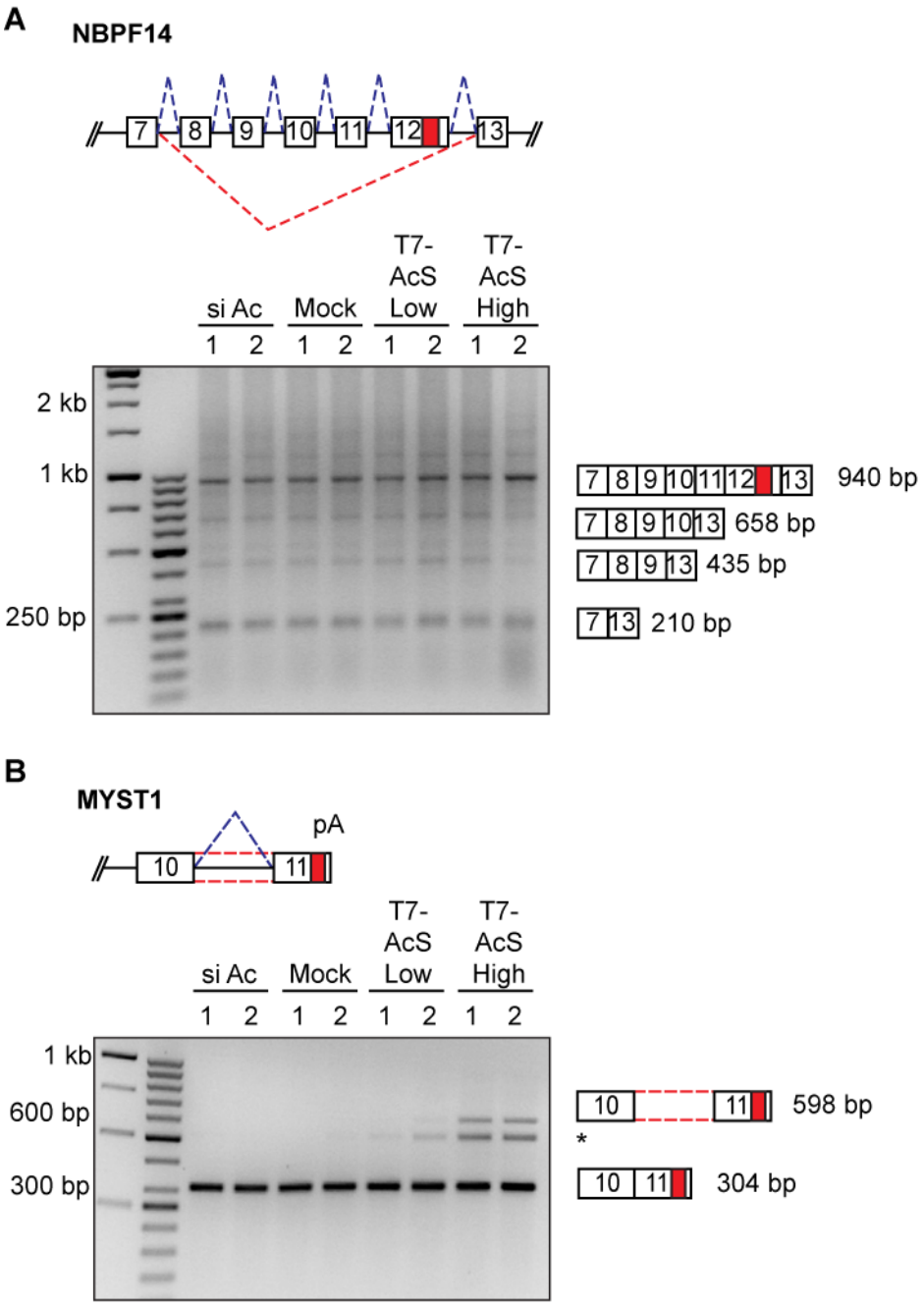


Figure 4.6. Alternative splicing regulation of Acinus target transcripts. The role of Acinus in the alternative splicing regulation of six of the transcripts identified by CLIP, where the binding site of Acinus is located in or near an alternative splicing event, were investigated. Products were amplified by RT-PCR (SuperScript™ III one-step RT-PCR system with Platinum® Taq, Invitrogen) from 500ng of total RNA from Acinus depleted (siAc), mock, or Acinus overexpressing (T7-AcS, Low and High) cells, and analysed on a 1.5% TBE agarose gel stained with ethidium bromide. A schematic of the alternative splicing event is shown above the gel, the red box indicating the location of the CLIP tag. Schematics and sizes of the potential products are shown to the right of the gel. See Materials and Methods for primers used. Two biological replicates are shown for each condition. * indicates a band of unconfirmed exon composition.

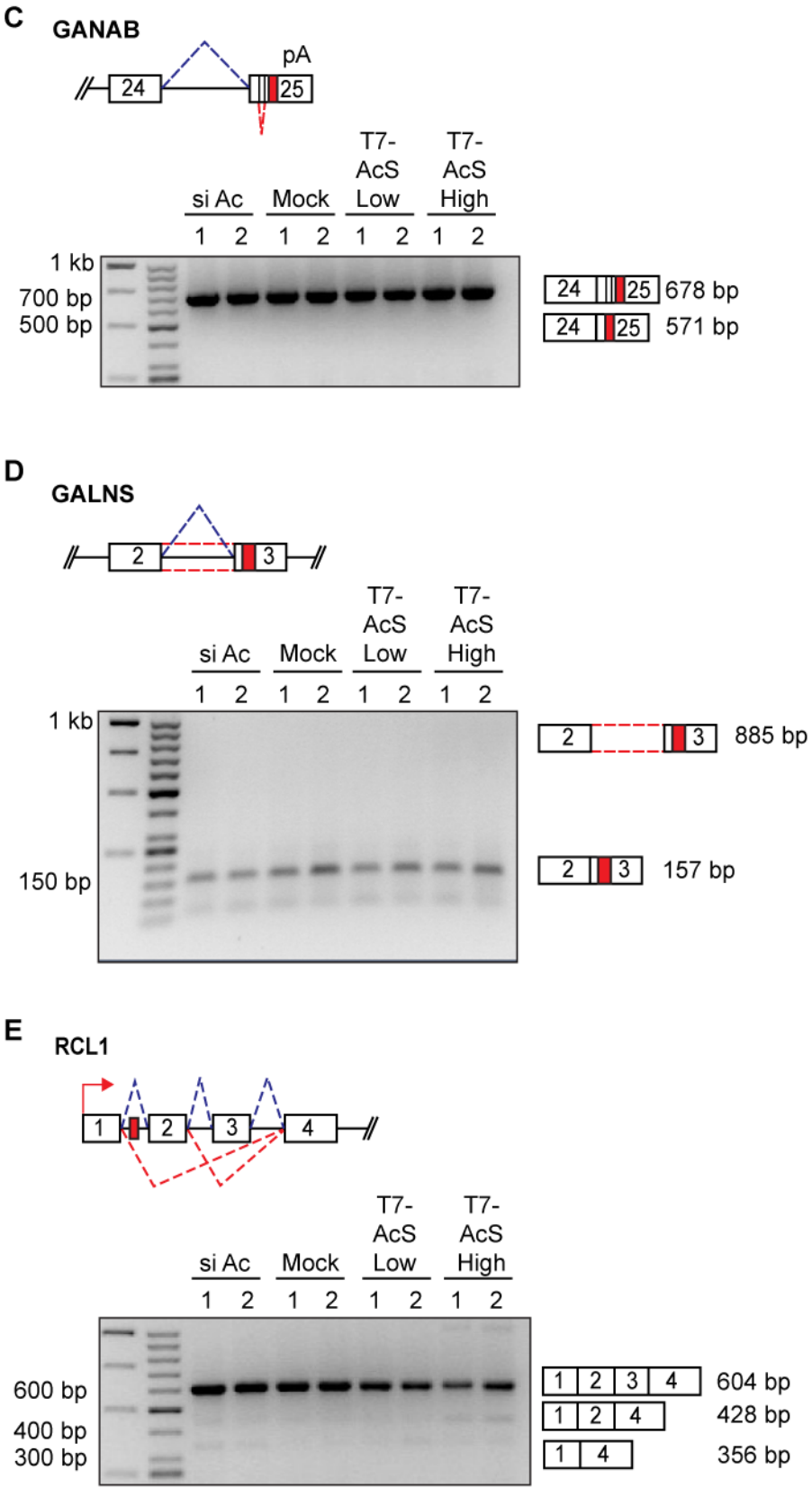


Figure 4.6. Alternative splicing regulation of Acinus target transcripts (continued).

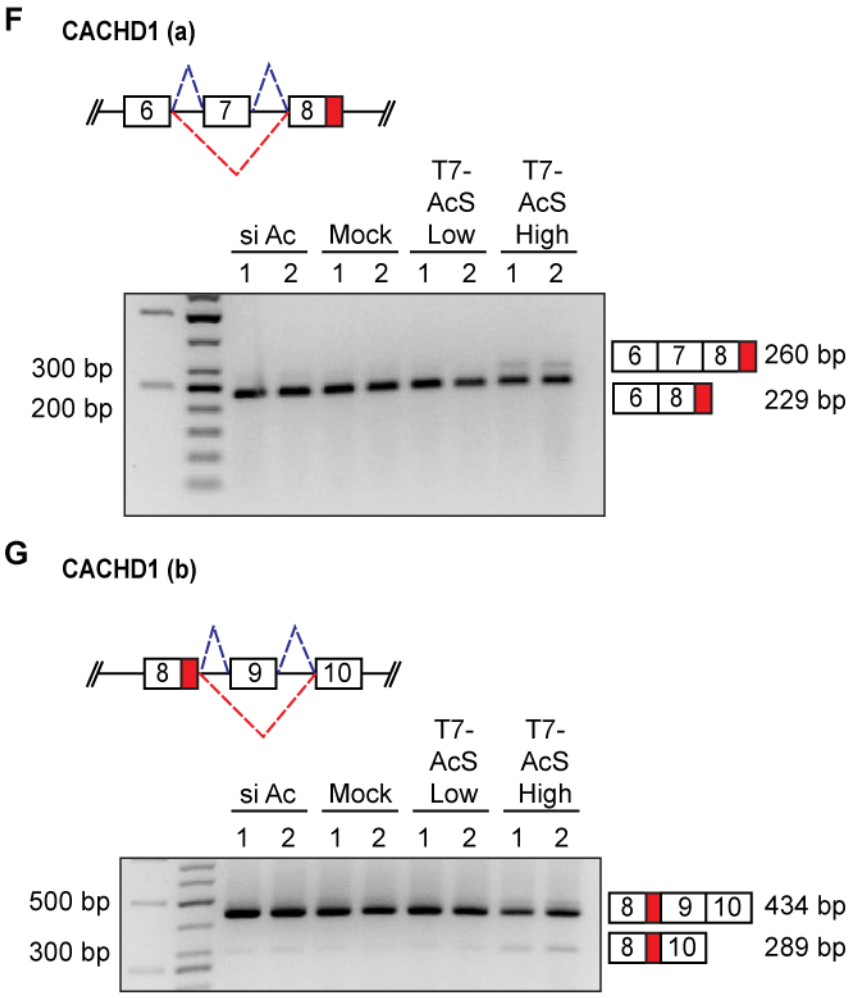


Figure 4.6. Alternative splicing regulation of Acinus target transcripts (continued).

4.4. Discussion

This chapter describes the identification of 40 protein-coding transcripts bound by Acinus, by use of the CLIP technique. No consensus binding sequence could be found amongst these sequences, although the extensive digestion of bound sequences when treated with the RNase A/T1 cocktail, as opposed to treatment with RNase T1 alone, suggests Acinus preferentially binds to pyrimidine rich sequences.

Acinus appears to bind almost exclusively to exonic regions. Two transcripts were identified where the binding site mapped to intronic regions, RCL1 and BPHL. However, it is possible these regions are unannotated exons, and further analysis of the sequence within the RCL1 transcript reveals a region of high conservation across species, suggesting the presence of a conserved intronic element or an unannotated exon. Therefore although Acinus may bind RNA in a nonsequence-specific manner, it does bind RNA in a location-specific manner.

The EJC, of which Acinus is a component, also binds RNA in a sequence-independent fashion but at a precise exonic location. Therefore it was hypothesised that Acinus may be a nuclear RNA-binding component of the EJC. However, positional analyses of the CLIP tags in relation to the exon-exon junction revealed only 31% coincided with the site of EJC deposition (Table 4.10), implying Acinus is unlikely to be a core RNA binding protein of the EJC.

Of the sequences identified as bound by Acinus 28% are located in, or adjacent to an alternative splicing event (Figure 4.3). Several of these alternative splicing events alter the primary structure of the encoded polypeptide (Table 4.8). Use of the proximal 5'ss in exon 4 of the SFRS10 transcript, encoding the splicing factor hTra2β, causes an in-frame deletion of 21 amino acids in part of the RS domain. As previously discussed, the RS domain is involved in pre-mRNA splicing, can target SR proteins to nuclear speckles and can direct nucleocytoplasmic shuttling. Therefore deletion of part of this domain may regulate the function of this protein in RNA processing.

Intron 10 retention of the MYST1 transcript results in the production of a protein isoform with a varying C-terminus to the canonical isoform. There are no identified key domains within this region, and the biological function of this isoform is

unknown. However, the alternative splicing event within this transcript was shown to be regulated by Acinus (Figure 4.6.B). Acinus promotes intron 10 retention of the MYST1 transcript in a concentration-dependent manner. Furthermore, MYST1 is one of the Acinus-bound transcripts that has been reported to act in apoptosis.

MYST1 (also known as hMOF) is a component of the multisubunit human histone acetyltransferase complex (hMSL) responsible for acetylation of histone H4 at lysine 16. Reduction in the levels of hMSL, and consequently reduced levels of H4K16 acetylation, are correlated with reduced transcription of some genes and with a G2/M cell cycle arrest (Smith *et al.*, 2005). MYST1 is ubiquitously expressed (Thomas *et al.*, 2007) in a similar pattern to that observed for Acinus (Figure 3.5), and has been shown to be essential for progression of mouse embryonic development past the blastocyst stage (Thomas *et al.*, 2008). Furthermore, MYST1 mutant embryos show abnormal chromatin morphology involving chromatin clumping before undergoing apoptosis. MYST1 also acetylates p53, an important modification required to induce *Bax* and *Puma* gene expression, thereby directing apoptosis (Sykes *et al.*, 2006).

MZF1 is another of the Acinus-bound transcripts that acts in apoptosis. It has been shown to interact with LDOC1, a leucine-zipper protein, and enhance its apoptotic inducing activity (Inoue *et al.*, 2005), while over-expression of MZF1 inhibits apoptosis (Hromas *et al.*, 1996, Robertson *et al.*, 1998).

Binding of Acinus to these transcripts may form part of the regulatory network involved in apoptosis regulation. In addition some of the transcripts identified as binding Acinus are involved in pre-mRNA splicing. These include the SR proteins, SFRS15 and SFRS10 (hTra2β), and UBE2I, a sumo-conjugating enxyme which localises to nuclear speckles (Ihara *et al.*, 2008). A number of the identified targets of Acinus are also involved in the cell cycle (ANAPC1, F2R and GNL3). Cell cycle regulation is an important aspect in cell proliferation and apoptosis, and it has previously been shown that phosphorylation of Acinus by SRPK2 results in upregulation of cyclin A1, while ablation of SRPK2 or Acinus arrest the cell cycle at G1 phase (Jang *et al.*, 2008).

Therefore Acinus is potentially involved in the regulation of pre-mRNA splicing, cell proliferation and apoptosis, and binds transcripts encoding proteins involved in all these processes.

However, the identification of the transcripts described in this chapter relied upon the over-expression of Acinus S. Therefore the detected protein-RNA interactions could be over-expression induced artefacts. This highlights the importance of validation of these targets by methods such as IP RT-PCRs, or gel shift analyses, which due to time constraints were not completed here.

Acinus has been described to act in the ASAP complex (Schwerk *et al.*, 2003) and addition of this complex to *in vitro* splicing reactions results in inhibition. However, the role of Acinus in splicing has not been previously described. Here, I show that Acinus acts as an alternative splicing regulator of both MYST1 and CACHD1 transcripts, identified as Acinus bound sequences by CLIP. Acinus may achieve this, in a similar manner to other SR proteins, by binding exonic elements to promote splicing and exon inclusion. However, further characterisation of how Acinus regulates these splicing events is required.

Chapter 5: HnRNP A1 CLIP

5.1. Introduction

The previous chapter describes the identification of about 40 transcripts bound by Acinus, using the CLIP technique. To complement this approach I also performed CLIP with the RNA processing factor hnRNP A1. Furthermore, analysis of hnRNP A1 targets may elucidate its function in the stress response.

The role of hnRNP A1 in alternative splicing has been extensively characterized. This protein predominantly binds to splicing silencers to antagonise the action of SR proteins (Figure 1.8B). However, hnRNP A1 also functions in processes such as nucleocytoplasmic transport and mRNA stability as discussed in Chapter 1.

The nucleocytoplasmic shuttling of hnRNP A1 has been shown to be subject to modulation by the mitogen-activated protein kinase kinase(3/6)-p38 ($MKK_{3/6}$-p38) signalling pathway in response to stress, resulting in its cytoplasmic accumulation (van der Houven van Oordt *et al.*, 2000). This was further shown to be concomitant with the phosphorylation of hnRNP A1, and the stress-induced phosphorylation sites in hnRNP A1 were mapped to a stretch of serines located in the F-peptide, adjacent to the M9 motif (Figure 1.7) (Allemand *et al.*, 2005). The transport receptors Trn 1 and Trn 2b interact directly with the M9 motif to mediate hnRNP A1 nuclear import (Pollard *et al.*, 1996; Rebane *et al.*, 2004). Phosphorylation of the F-peptide abrogates interactions between hnRNP A1 and transportin resulting in its cytoplasmic accumulation (Allemand *et al.*, 2005).

The Mnk1/2 protein kinases, which act downstream of p38, are responsible for the stress-induced phosphorylation of hnRNP A1, which result in its accumulation in cytoplasmic foci known as stress granules (Guil *et al.*, 2006). Furthermore, depletion of hnRNP A1 affects the recovery of cells from stress, suggesting a physiologically significant role for hnRNP A1 in the stress response, which may involve the regulation of key transcripts, as redistribution of hnRNP A1 to the cytoplasm results in changes in alternative splicing patterns (van der Houven van Oordt *et al.*, 2000).

Again, CLIP was used to identify the *in vivo* RNA targets of hnRNP A1, which may assist in understanding how it regulates gene expression. One aim was to determine if these targets changed during the stress response, and if hnRNP A1 regulated the splicing of these targets in a stress-dependent manner. It was also

hoped that by performing CLIP on two RNA-binding proteins with different cellular roles, it could be determined if they bind distinct or overlapping sets of RNAs, or if they exhibit distinct binding site preferences, for example preferential binding to intronic or exonic regions.

5.2. Results

HnRNP A1 CLIP

The conditions for hnRNP A1 CLIP had previously been determined (Sonia Guil, personal communication, see Materials and Methods for details). CLIP was performed with endogenous hnRNP A1 from HeLa cells. Figure 5.1 shows UV-dependent, RNase-sensitive complexes immunoprecipitated by the anti-hnRNP A1 monoclonal antibody. The RNA was extracted from the membrane and amplified by RT-PCR. BLAST was used to align sequences to the human genome allowing for two mis-matches or gaps within the alignment.

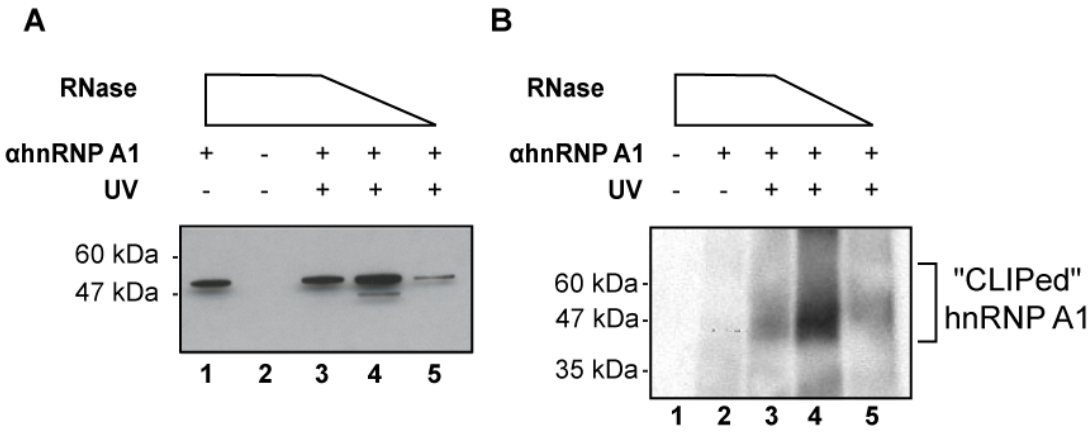
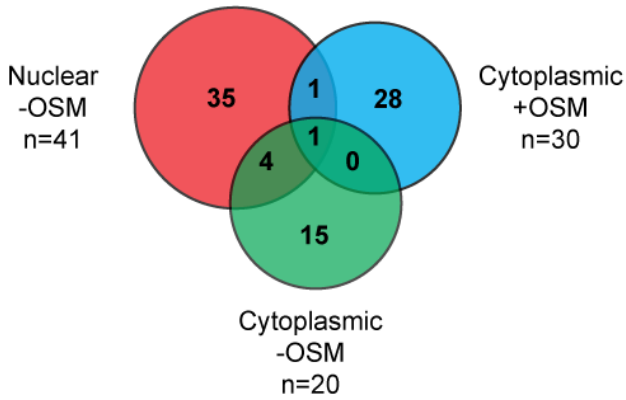


Figure 5.1. Analysis of hnRNP A1-RNA complexes by CLIP. Complexes were isolated with anti-hnRNP A1 (4B10, ImmunoQuest) immunoprecipitation from the nuclear fraction of non-stressed HeLa cells, however, the data is representative of all experimental conditions (nuclear, cytoplasmic ± osmotic shock). (**A**) Western blot analysis of the immunoprecipitation reaction to show the mobility of the proteins not bound to RNA. (**B**) Analysis of SDS-PAGE separated protein-RNA complexes after RNA digestion and T4 PNK assay.

CLIP with hnRNP A1 was performed under a variety of experimental conditions designed to elucidate the role of hnRNP A1 in the stress response. HeLa cells, in the absence or presence of osmotic shock, were fractionated into cytoplasmic and nuclear extracts. Targets bound to hnRNP A1 under these conditions were then determined by the CLIP technique. A summary of the number of transcripts identified in each condition, and any overlapping transcripts, is shown in Figure 5.2. However, due to the small sample sizes obtained in each condition the data is presented together as a whole in this chapter.

A



B

Nuclear –OSM	Nuclear +OSM	Cytoplasmic –OSM	Cytoplasmic +OSM
MIRH1	MIRH1	MIRH1	MIRH1
U12snRNA		U12snRNA	
FER		FER	
KIF1B		KIF1B	
ANP32A		ANP32A	
RN7SK			RN7SK
n=41	n=6	n=20	n=30

Figure 5.2. Analysis of transcripts bound by hnRNP A1 with respect to their experimental identification. CLIP was performed with nuclear and cytoplasmic cellular fractions in the presence and absence of osmotic shock (OSM). (**A**) Venn diagram showing the numbers of transcripts identified in each condition. (**B**) Table detailing the transcripts which were found in more than one condition. See Appendix 2 for further information on the CLIP tags.

Identification of hnRNP A1 targets

To determine if the endogenous binding sites of hnRNP A1 can provide more information regarding the role of hnRNP A1 in alternative splicing, the transcripts targeted by hnRNP A1 were manually curated for alternative splicing events using Ensembl, UCSC genome browser and Fast-DB databases.

Of the 89 binding sites present in protein-coding genes, 30 (34%) of these were located in, or adjacent to, an alternative splicing event (Tables 5.1-5.6, Figure 5.3). These include hnRNP A1 binding sites located in alternative cassette exons (SPTAN1 and TMEM189/UBE2V1), in an exon with an alternative 3' splice site (SHROOM3), in exons with alternative 5' splice sites (CDK4, HNRNPA1, RPS16 and SHROOM3), and in exons with retained introns (CRKL and MIRH1) or internal exon deletions (BASP1 and EFTUD2). HnRNP A1 binding sites were also found adjacent to alternative cassette exons, in both exonic regions (CDK4, HMGCR, HN1, HNRNPA1, LPHN2, SHROOM3 and SULT4A1, Table 5.5) and intronic regions (CDKAL1, COG5, FER, IL6R, KIF1B, JMJD1C, MID1, SLC25A43, UBR2, UNC45B and XPO6, Table 5.6). The presence of hnRNP A1 binding sites in close proximity to these alternative splicing events could indicate a role for hnRNP A1 in the splicing regulation of these transcripts.

Several of the alternative splicing events lead to PTCs and are therefore potential substrates for NMD (SHROOM3, CDK4, EFTUD2, HN1, LPHN2, SLC25A43, XPO6, MID1 and UBR2). Other events alter the primary structure of the encoded polypeptide, which in some cases may affect the biological function of the protein (TMEM189/UBE2V1, SPTAN1, RPS16, HNRNPA1, HMGCR, IL6R, KIF1B and UBR2, Table 5.7).

The remainder of the transcripts identified by CLIP as bound to hnRNP A1 are listed in Table 5.8. See Appendix 2 for full sequence details of the CLIP tag, genomic co-ordinates and Ensembl gene ID.

Gene ID	Exon Position	Gene Description	Localisation	Function
SPTAN1	21	Spectrin, alpha, non-erythrocytic 1 (alpha-fodrin)	Nuclear	Structural constituent of cytoskeleton
TMEM189/ UBE2V1	6	Transmembrane protein 189/ ubiquitin-conjugating enzyme E2 variant 1	Cytoplasmic	Transcription regulation; Protein polyubiquitination

Table 5.1. Alternative cassette exons bound by hnRNP A1. Exons were annotated manually according to the FastDB database (www.fast-db.com, de la Grange *et al.*, 2005). Column headers: Gene ID, official gene symbol; Exon/Intron Position, number of exon/intron within the primary transcript; Gene Description; Localisation, subcellular fraction used for CLIP experiment; Function, annotated function of the protein product.

Gene ID	Exon Position	Gene Description	Localisation	Function
SHROOM3	7	Shroom family member 3	Cytoplasmic	Cell morphogenesis

Table 5.2. HnRNP A1 target exons containing alternative 3' splice sites.

Gene ID	Exon Position	Gene Description	Localisation	Function
CDK4	3	Cyclin-dependent kinase 4	Cytoplasmic	Cell cycle regulation
HNRNPA1	6	Heterogeneous nuclear ribonucleoprotein A1	Cytoplasmic	mRNA processing
RPS16	2	Ribosomal protein S16	Cytoplasmic	Translation
SHROOM3	7	Shroom family member 3	Cytoplasmic	Cell morphogenesis

Table 5.3. HnRNP A1 target exons containing alternative 5' splice sites.

Gene ID	Exon Position	Gene Description	Localisation	Function
BASP1	2	Brain abundant, membrane attached signal protein 1	Cytoplasmic	Cytoskeletal component
CRKL	3	v-crk sarcoma virus CT10 oncogene homolog (avian)-like	Nuclear	Protein phosphorylation; Intracellular signalling
EFTUD2	3	Elongation factor Tu GTP binding domain containing 2	Nuclear	mRNA processing
MIRH1	3	MicroRNA host gene (non-protein coding) 1	Nuclear; Cytoplasmic	Encodes micro RNAs

Table 5.4. HnRNP A1 target exons containing a retained intron.

Gene ID	Exon Position	Gene Description	Localisation	Function
CDK4	3	Cyclin-dependent kinase 4	Cytoplasmic	Cell cycle regulation
HMGCR	12	3-hydroxy-3-methylglutaryl-Coenzyme A reductase	Cytoplasmic	Germ cell migration
HN1	1,4	Hematological and neurological expressed 1	Nuclear	Nervous system development
HNRNPA1	6	Heterogeneous nuclear ribonucleoprotein A1	Cytoplasmic	mRNA processing
LPHN2	23	Latrophilin 2	Nuclear	G-protein coupled receptor activity
SHROOM3	7	Shroom family member 3	Cytoplasmic	Cell morphogenesis
SULT4A1	8	Sulfotransferase family 4A, member 1	Nuclear	Sulfotransferase activity

Table 5.5. HnRNP A1 target exons located adjacent to alternative cassette exons.

Gene ID	Intron Position	Gene Description	Localisation	Function
CDKAL1	5	CDK5 regulatory subunit associated protein 1-like 1	Nuclear	Unknown
COG5	17	Component of oligomeric golgi complex 5	Nuclear	Intra-golgi vesicle mediated transport
FER	4	Fer (fps/fes related) tyrosine kinase	Nuclear; Cytoplasmic	Intracellular signalling
IL6R	7	Interleukin 6 receptor	Nuclear; Cytoplasmic	Cell proliferation; Signal transduction; Immune response
KIF1B	24	Kinesin family member 1B	Nuclear; Cytoplasmic	Microtubule based movement; Neuromuscular synaptic transmission
JMJD1C	4	Jumonji domain containing 1C	Cytoplasmic	Transcription regulation
MID1	8	Midline 1 (Opitz/BBB syndrome)	Nuclear	Microtubule cytoskeleton organization and biogenesis
SLC25A43	3	Solute carrier family 25, member 43	Nuclear	Mitochondrial transport
UBR2	12	Ubiquitin protein ligase E3 component n-recognin 2	Nuclear	Ubiquitination cycle
UNC45B	4	Unc-45 homolog B (C. elegans)	Cytoplasmic	Skeletal muscle development
XPO6	2	Exportin 6	Nuclear	Nucleocytoplasmic transport

Table 5.6. HnRNP A1 target introns located adjacent to alternative cassette exons.

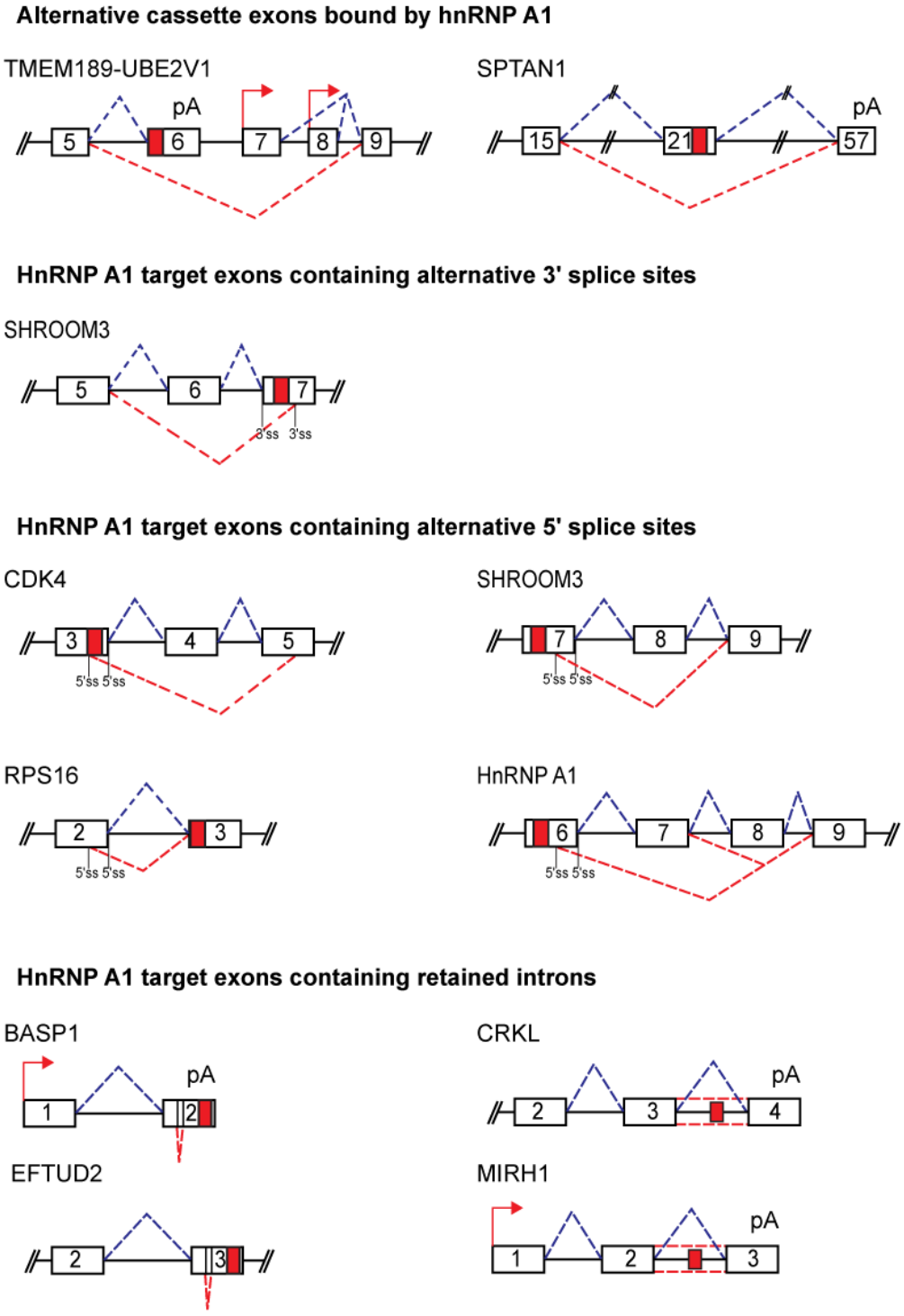


Figure 5.3. RNA targets of hnRNP A1 identified by CLIP which are located in or nearby sites of alternative splicing events. Red box indicates CLIP tag location within the transcript. Splicing profiles determined from FastDB (www.fast-db.com, de la Grange *et al.*, 2005) or UCSC (http://genome.ucsc.edu, Kent *et al.*, 2002) databases. pA: polyadenylation site; ss: splice site; red arrow: transcription start site.

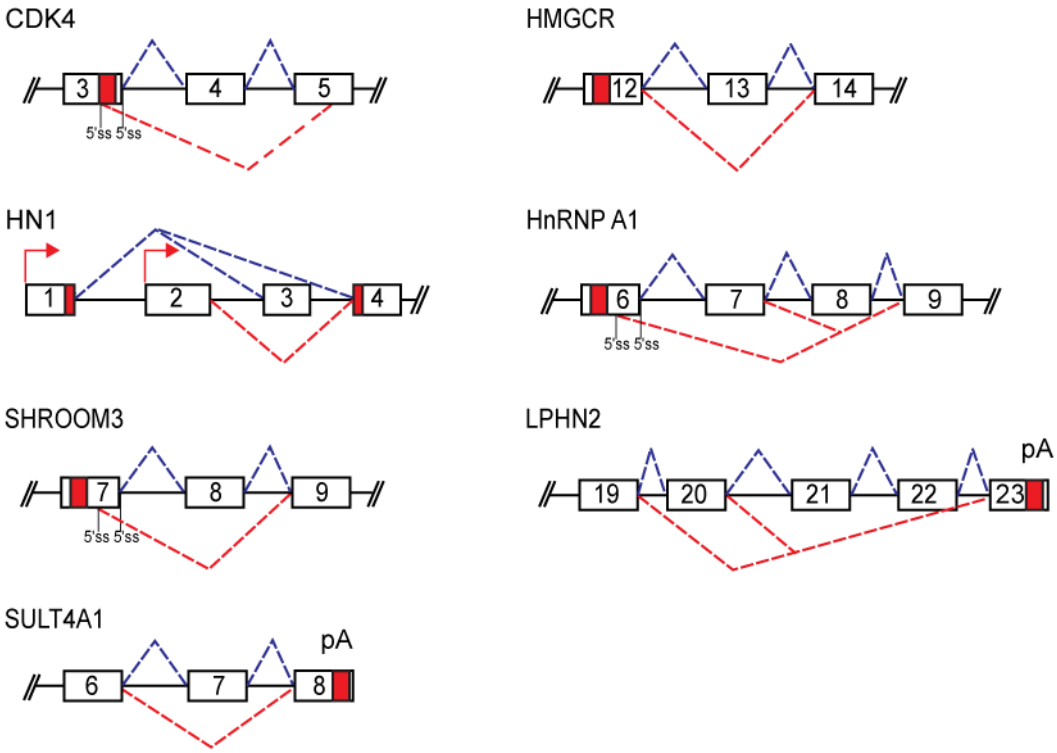


Figure 5.3. RNA targets of hnRNP A1 identified by CLIP which are located in or nearby sites of alternative splicing events (continued).

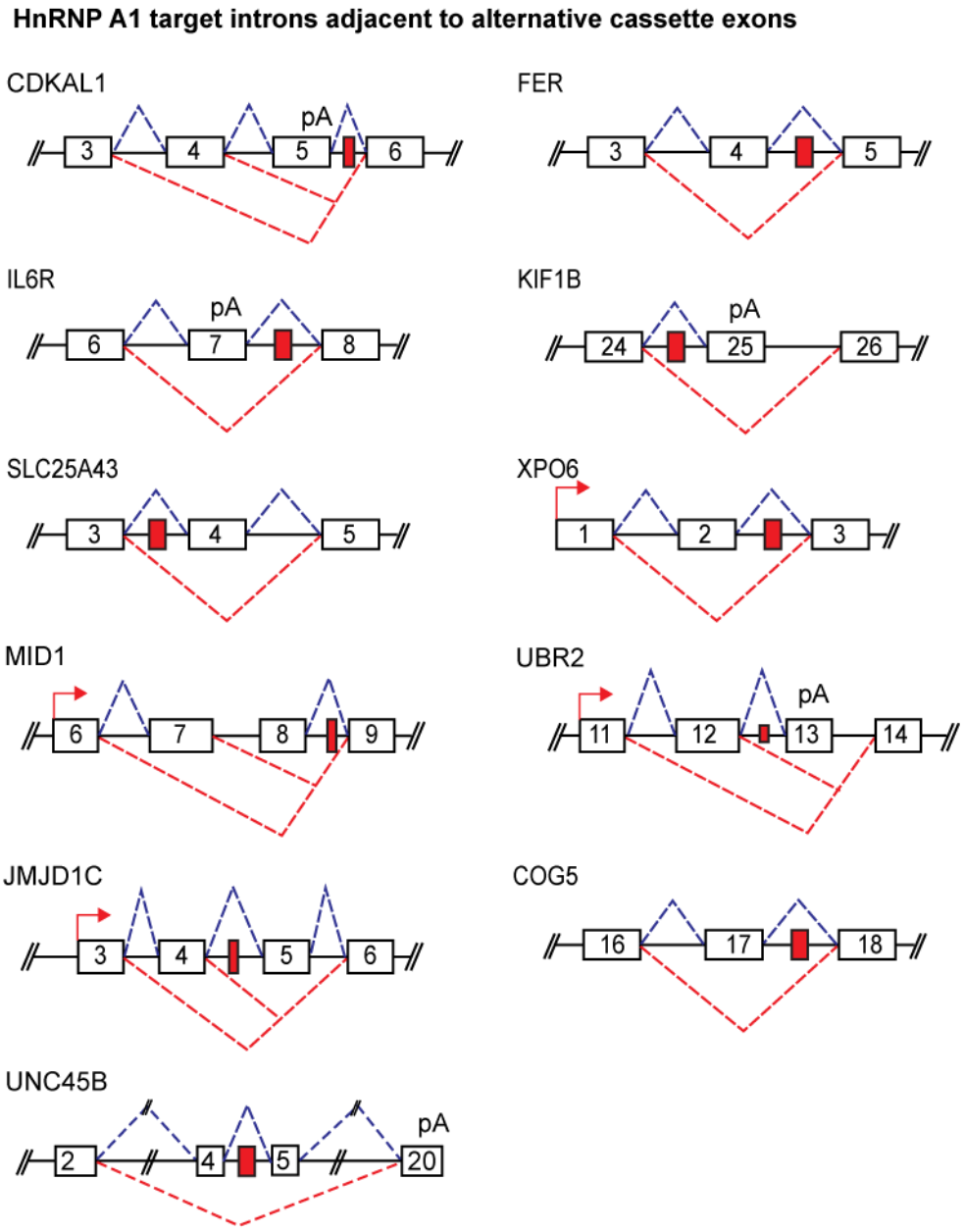


Figure 5.3. RNA targets of hnRNP A1 identified by CLIP which are located in or nearby sites of alternative splicing events (continued).

Transcript	Alternative splicing event	Effect on protein product
BASP1	Internal exon deletion in exon 2	Event occurs in 3' UTR - likely to induce NMD
CDK4	Alternative 5' ss in exon 3	Use of proximal 5' ss introduces a PTC - possible substrate for NMD
CRKL	Intron 3 retention	Event occurs in 3' UTR - likely to induce NMD
EFTUD2	Internal exon deletion in exon 3	Introduces a PTC - possible substrate for NMD
FER	Exon 4 skipping	Event occurs in 5' UTR
HMGCR	Exon 13 skipping	In-frame deletion of 53 amino acids within the catalytic domain
HN1	Alternate promoter in exon 1 and exon 3 inclusion	Introduces a PTC - possible substrate for NMD
HN1	Alternate promoter in exon 2 and exon 3 skipping	Event occurs in 5' UTR
HNRNPA1	Exon 8 inclusion	Inclusion of exon 8 (7B) results in the production of hnRNP A1B isoform
HNRNPA1	Alternative 5' ss in exon 6 and exons 7-8 skipping	In-frame deletion of 48 amino acids resulting in loss of the RGG-box motif
IL6R	Exon 7 skipping	Two protein isofoms are produced; if exon 7 is included it introduces a stop codon and a short, soluble protein is produced, while exon 7 skipping results in the production of a long isoform encoding a transmembrane protein.
JMJD1C	Exons 4 and/or 5 skipping	Event occurs in 5' UTR.
KIF1B	Exon 25 skipping	Two protein isoforms are produced; if exon 25 is included it introduces a stop codon and a short isoform, while exon 25 skipping resuls in the production of a long isoform
LPHN2	Exons 20-22 skipping	Inclusion of exons 21 and 22, or skipping of exons 20-22 introduces a PTC - possible substrates for NMD
MID1	Exons 7 and/or 8 skipping	Introduces PTCs - possible substrates for NMD
RPS16	Alternative 5' ss in exon 2	Use of proximal 5' ss causes an in-frame deletion of 17 amino acids
SHROOM3	Exon 6 skipping, alternative 3' ss and 5' ss of exon 7, exon 8 skipping	Introduces a PTC - possible substrate for NMD
SPTAN1	Exons 16-56 skipping	Production of a truncated protein lacking 15 (of 23) spectrin repeats
TMEM189/ UBE2V1	Exons 6-8 skipping Alternate promoters in exon 7 and 8	Production of various isoforms with differing C-termini. Proteins produced from transcripts starting in exons 7 or 8 lack the ubiquitin-conjugating E2 domain
UBR2	Exons 12 and/or 13 skipping	In-frame deletion of 32 amino acids as a result of exon 12 skipping. Inclusion of exon 13 introduces a stop codon and production of a shorter isoform.
XPO6	Exon 2 inclusion	Introduces a PTC - possible substrate for NMD

Table 5.7. Effect on protein product of alternative splicing of hnRNP A1 bound transcripts. The effect of the described alternative splicing event was determined by manually manipulating sequences and by searching databases for the potential protein product isoforms.

Gene ID	CLIP Position	Gene Description	Localisation	Function
ABCF2	Exon 3, 4	ATP-binding cassette, sub-family F, member 2	Cytoplasmic	ATPase activity
ANP32A	Exon 7, Intron 1	Acidic leucine-rich nuclear phosphoprotein 32 family member A	Nuclear	Nucleocytoplasmic transport
ANP32B	Exon 3	Acidic leucine-rich nuclear phosphoprotein 32 family member B	Cytoplasmic	Unknown
ARHGAP15	Intron 6	Rho GTPase activating protein 15	Cytoplasmic	Signal transduction
ARID4A	Exon 24	AT rich interactive domain 4A	Cytoplasmic	Chromatin binding; Transcription regulation
ASB7	Exon 4	Ankyrin repeat and SOCS box protein 7	Cytoplasmic	Ubiquitination cycle
AYTL2	Intron 2	Acyltransferase-like 2	Cytoplasmic	Lipid metabolism
C11ORF47	Exon 4	Chromosome 11 open reading frame 47	Nuclear	Unknown
C16ORF52	Intron 1	Chromosome 16 open reading frame 52	Nuclear	Unknown
C2ORF27	Intron 1	Chromosome 2 open reading frame 27	Nuclear	Unknown
C6ORF103	Intron 1	Chromosome 6 open reading frame 103	Cytoplasmic	Unknown
CAPZA1	Intron 3	Capping protein (actin filament) muscle Z-line, alpha 1	Nuclear	Cell motility
CHD3	Exon 40	Chromodomain-helicase-DNA-binding protein 3	Cytoplasmic	Chromatin binding; Transcription regulation
CRIM1	Intron 1	Cysteine rich transmembrane BMP regulator 1 (chordin-like)	Nuclear	Proteolysis; Cell adhesion
DNAH17	Intron 19	Dynein, axonemal, heavy chain 17	Nuclear	Microtubule-based movement
DYRK1A	Exon 14	Dual specificity tyrosine-phosphorylation-regulated kinase 1A	Cytoplasmic	Protein phosphorylation
EEF1G	Exon 4	Eukaryotic translation elongation factor 1 gamma	Cytoplasmic	Translational elongation
FAM38A	Exon 31, 32	Family with sequence similarity 38, member A	Cytoplasmic	Unknown
FAM62B	Exon 11, 12	Family with sequence similarity 62 (C2 domain containing) member B	Nuclear	Sodium ion transport
FARP1	Intron 2	FERM, RhoGEF (ARHGEF) and pleckstrin domain protein 1	Cytoplasmic	Glycolysis; Signal transduction
FLJ23584	Exon 1	Hypothetical FLJ23584	Nuclear	Unknown
FNDC8	Intron 3	Fibronectin type III domain-containing protein 8	Cytoplasmic	Unknown
FUSIP1	Exon 6	FUS interacting protein (serine-arginine rich) 1	Nuclear	mRNA processing
GABARAPL1	Exon 7	GABA(A) receptor-associated protein like 1	Cytoplasmic	GABA receptor binding
GEMIN8	Exon 5	Gem-associated protein 8	Nuclear	mRNA processing
GRIN2A	Intron 4	Glutamate receptor, ionotropic, N-methyl D-aspartate 2A	Nuclear	Ion transport
HBG1	Exon 3	Haemoglobin, gamma A	Cytoplasmic	Oxygen transport
KCTD7	Intron 3	Potassium channel tetramerisation domain containing 7	Nuclear	Ion transport
KIF1B	Intron 24	Kinesin family member 1B	Nuclear	Microtubule-based movement

Table 5.8. Other hnRNP A1 target transcripts.

Gene ID	CLIP Position	Gene Description	Localisation	Function
LYN	Intron 1	Yamaguchi sarcoma viral (v-yes-1) related oncogene homolog	Cytoplasmic	Protein phosphorylation
MACROD2	Intron 3	MACRO domain containing 2	Nuclear	Unknown
MAT2A	Exon 9	Methionine adenosyltransferase II, alpha	Cytoplasmic	Methionine adenosyltransferase activity
MESDC2	Intron 3	Mesoderm development candidate 2	Nuclear	Mesoderm development
MPP3	Intron 10	Membrane protein, palmitoylated 3 (MAGUK p55 subfamily member 3)	Nuclear	Signal transduction
NBPF14	Exon 3	Neuroblastoma breakpoint family member 14	Cytoplasmic	Unknown
NEAT2	Exon 1	Nuclear enriched abundant transcript 2	Cytoplasmic	Unknown
NKTR	Intron 2	Natural killer-tumor recognition sequence	Nuclear	Isomerase activity
NOS1AP	Intron 2	Nitric oxide synthase 1 (neuronal) adaptor protein	Nuclear	Nitric oxide synthesis regulation
PCDH7	Intron 2	Protocadherin 7	Nuclear	Calcium ion binding
PDS5A	Exon 19	PDS5, regulator of cohesion maintenance, homolog A (S. cerevisiae)	Cytoplasmic	Cell cycle regulation
PEX13	Intron 1	Peroxisome biogenesis factor 13	Cytoplasmic	Protein transport
RABEP1	Intron 1	Rabaptin, RAB GTPase binding effector protein 1	Nuclear	Endocytosis
RAPTOR	Intron 1	Regulatory-associated protein of mTOR	Cytoplasmic	Signal transduction
RMND5A	Intron 5	Required for meiotic nuclear division 5 homolog A (S. cerevisiae)	Nuclear	Unknown
RN7SK	Exon 1	RNA, 7SK, nuclear	Nuclear	Transcription regulation
RPS26	Intron 2	Ribosomal protein S26	Nuclear	Translation
RPS27	Exon 1	Ribosomal protein S27	Nuclear	Translation
RYR2	Intron 33	Ryanodine receptor 2 (cardiac)	Cytoplasmic	Ion transport
SFRS3	Intron 2	Splicing factor, arginine/serine-rich 3	Cytoplasmic	mRNA processing
SLC39A1	Exon 4	Solute carrier family 39 (zinc transporter), member 1	Cytoplasmic	Ion transport
SMAD2	Exon 3	SMAD family member 2	Cytoplasmic	Transcription regulation
SNORD3B-2	Exon 1	Small nucleolar RNA, C/D box 3B-2	Cytoplasmic	Ribosomal RNA modification
STK36	Exon 7,8	Serine/threonine-protein kinase 36	Cytoplasmic	Protein phosphorylation
TMEM49	Intron 11	Transmembrane protein 49	Cytoplasmic	Unknown
U12 snRNA	Exon 1	U12 small nuclear RNA	Nuclear; Cytoplasmic	mRNA processing
U1A snRNA	Exon 1	U1A small nuclear RNA	Nuclear	mRNA processing
U6 snRNA	Exon 1	U6 small nuclear RNA	Nuclear	mRNA processing
USP24	Exon 14	Ubiquitin specific peptidase 24	Cytoplasmic	Ubiquitination cycle
WAC	Exon 4	WW domain containing adaptor with coiled-coil	Nuclear	Unknown
ZC3HC1	Intron 5	Zinc finger, C3HC-type containing 1	Cytoplasmic	Protein ubiquitinylation
ZNRF3	Exon 10	Zinc and ring finger 3	Cytoplasmic	Ion binding

Table 5.8. Other hnRNP A1 target transcripts (continued).

Analysis of sites to which hnRNP A1 binds

A number of the sequences identified by CLIP as binding hnRNP A1, contain the hnRNP A1 SELEX winner sequence UAGGG(A/U) (Burd and Dreyfuss, 1994; Figure 5.4). Furthermore, 45% of CLIP tags contain the core UAG binding motif.

RNA tag	Transcribed sequence
ARID4A	CACUUUCACUAAGAUGUAUUUGAACACUUGGUGAGUAGGG
ASB7	ACUCAGCUAGGAUGUUACACCAUCAUUGUCGAAGGAACCCUGAGCUCCUG
TMEM189/ UBE2V1	CAUCACCACAGGCUGGCUCAACUACCCUCUGGAGAAGAUAGGCUUCUG
BASP1	CCUGAGGAGAGUGGAGGACGUCAUCUGGAUAGGAUCUCAAUGCCAAUCCUCCAUUCUUCCUCUCCAG
HN1	CAACAGCAGGAAUAGCUCCCGAGUUUUGCGGCCUCCAGGUGGUGGAUCCAAUUUUUCAUUAGGUUUUGAUG
LPHN2	AUCAGCUUCAGAUGUGCUACCAGAUCAGCAGGGGCAAUAGUGAUGGUUAUAUAAUCCCCAUU
MESDC2	UCAGUCAUGAGAGGCAAAGUGGAAGGUGGGUUGGGCCAAACACAGGGUCAGCUUUAGGGCC
MIRH1	UGCCAGAAGGAGCACUUAGGGCAGUAGAUGCUAAUCU
FAM62B	CAAGGAGAACCUCAGUCCAAANUGGAAUGAAGUCUAUGAGGCUUUAGUGUAUGAACAUCCUGGAC

Figure 5.4. HnRNP A1 CLIP tags which contain sequences similar to the SELEX winner sequence UAGGGA/U (Burd and Dreyfuss, 1994).

The motif finding algorithm MEME was used to search nuclear, cytoplasmic, exonic or intronic CLIP tags for over-represented sequences. A purine-rich consensus motif is present within nuclear sequences (containing both intronic and exonic sequences) targeted by hnRNP A1, with a highly significant e-value of 4.1 x 10^{-5} (Figure 5.5.A). A pyrimidine-rich consensus motif was identified within the cytoplasmic (containing both intronic and exonic sequences) CLIP tags, although the e-value of 4.4 x 10^{-1} is less significant (Figure 5.5.B). MEME was unable to identify a consensus motif with significant e-value within the exonic or intronic binding sequences (from the combined nuclear and cytoplasmic sequences).

The purine-rich consensus motif, (A/G)AGUGA, identified from the nuclear CLIP tags, is similar although not equivalent to the binding sequence identified by SELEX.

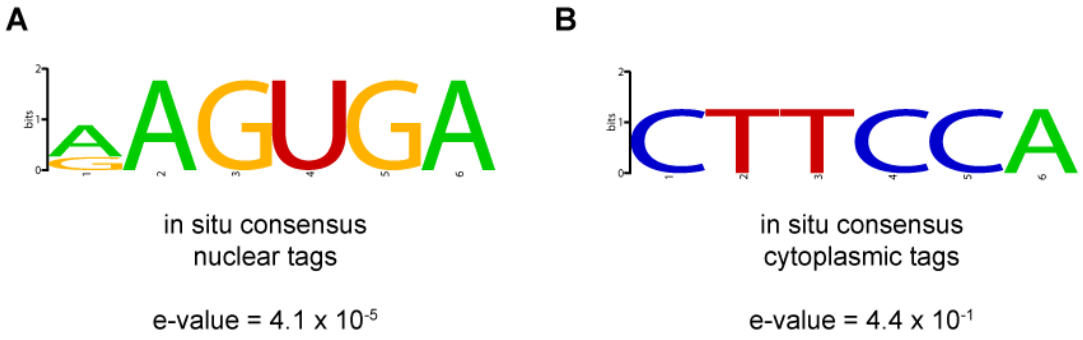


Figure 5.5. Identification of the hnRNP A1 binding site. The nuclear (**A**), cytoplasmic (**B**), exonic or intronic CLIP tags were analysed using MEME (http://meme.sdsc.edu/meme, Bailey and Elkan, 1994) to identify sequence patterns that are over-represented. The e-value is a parameter that describes the number of motifs one can expect to see by chance when searching a similarly sized set of random sequences. The lower the e-value, the more significant the match. MEME was unable to identify a consensus motif within the exonic or intronic binding sites. Sequence logos generated using WebLogo (http://weblogo.berkeley.edu, Crooks *et al.*, 2004).

HnRNP A1 is involved in the stress response

As discussed in the introduction, the predominantly nuclear hnRNP A1 protein accumulates in the cytoplasm in response to environmental stresses. The RNA binding proteins TIA-1 and TIAR associate with SGs in cells exposed to environmental stress, and are considered robust markers of SGs (Kedersha *et al.*, 1999). Upon OSM endogenous hnRNP A1 is found to be widely distributed throughout the cytoplasm, but concentrated in SGs as shown by co-localisation studies with TIA-1 (Figure 5.6).

Osmotic stress does not appear to affect the levels of endogenous hnRNP A1 protein (Figure 5.7.B) or T7-tagged hnRNP A1 protein (Figure 5.7.C), as assayed by Western blot. Neither does osmotic stress affect the level of hnRNP A1 depletion by RNAi (Figure 5.6.A).

There are a number of interesting questions regarding the role of hnRNP A1 in the stress response. It is possible that hnRNP A1 regulates the trafficking of RNAs to SGs, and affects their subsequent metabolism in times of cellular stress. To try to address these questions I aimed to determine if the splicing profiles of transcripts targeted by hnRNP A1, identified here by CLIP, changed in the presence or absence of stress, and if hnRNP A1 modulates these splicing profiles.

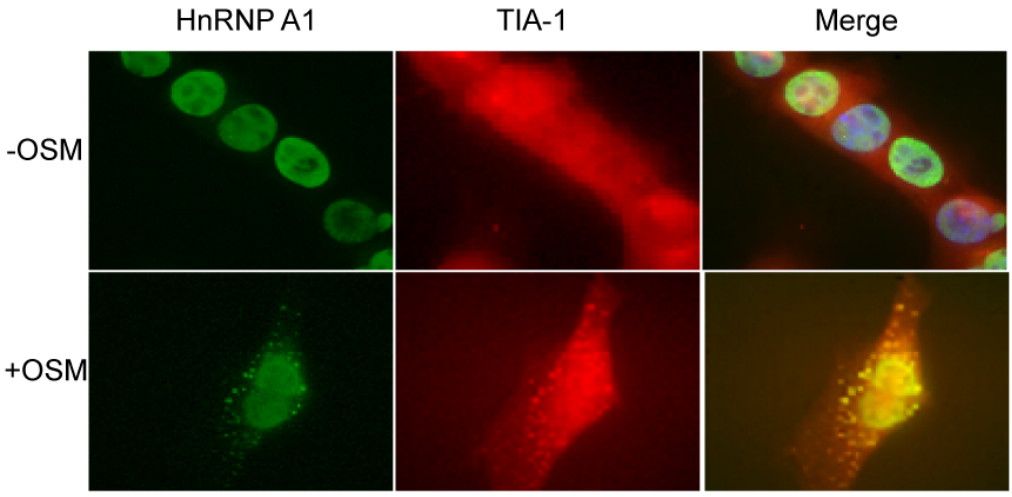


Figure 5.6. HnRNP A1 is recruited to stress granules (SGs) in response to osmotic shock (OSM). The subcellular localization of hnRNP A1 was determined by immunofluorescence using a monoclonal antibody against endogenous hnRNP A1 (4B10, Immunoquest) (left panels; green). Cells were exposed to OSM by treatment with 600 mM sorbitol for 2.5 h. SGs were visualised by counterstaining for TIA-1 (center panels; red). Merged images shown in right panels. (Data previously submitted for degree of Masters by Research, Edinburgh University, 2005).

HnRNP A1 modulates the alternative splicing of several transcripts

To determine if hnRNP A1 regulates the alternative splicing of the transcripts identified by CLIP, 16 of the transcripts where the binding site of hnRNP A1 is in, or nearby an alternative splicing event, were investigated. The levels of hnRNP A1 were manipulated in HeLa cells by RNAi depletion (Figure 5.7.A) or transient expression of T7-hnRNP A1 (Figure 5.7.C). Total RNA prepared from these cells was used in RT-PCR based assays, with primers designed in regions flanking the alternative processing event, to determine if the levels of hnRNP A1, or induction of the stress-signalling pathway, modulate the splicing pattern.

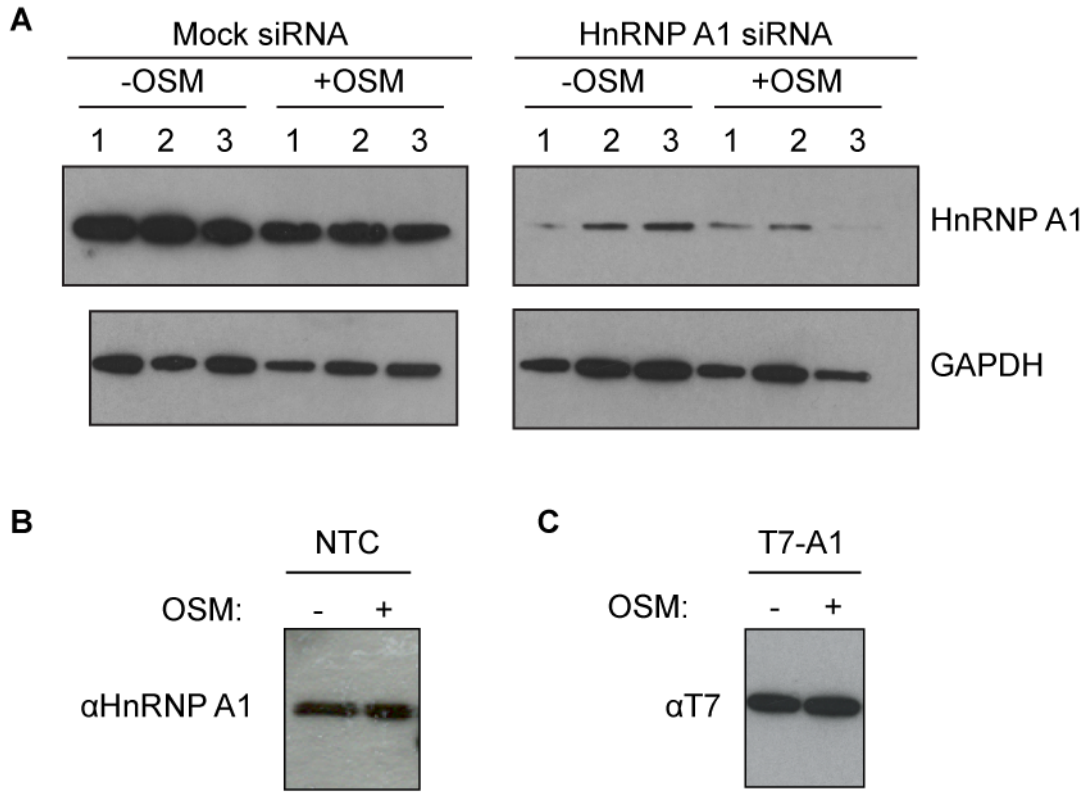


Figure 5.7. Analysis of hnRNP A1 knockdown and overexpression in the presence or absence of osmotic shock. (**A**) HeLa cells were transfected with siHnRNPA1 before being osmotically shocked (OSM) by the addition of 600 mM sorbitol for 2.5 h. Extracts were blotted with an anti-hnRNP A1 antibody (4B10, ImmunoQuest) and GAPDH was used as a loading control (ab9485, AbCam). OSM does not affect the level of hnRNP A1 depletion, nor does it affect endogenous levels of hnRNP A1 (**B**) or the overexpression of T7-tagged hnRNP A1 (T7-A1) (**C**). Three biological replicates are shown for each condition. NTC: non-transfected control.

As discussed in Chapter 4, it was essential to use RNA of high quality and integrity in these assays. Figure 5.8.A shows the electropherograms and densitometry plots for each RNA sample, while other information including the RIN is shown in Figure 5.8.B. The RNA samples were also used to amplify 1 kb fragments of both Beta-actin and GAPDH using intron-spanning primers to further assess the quality of the RNA (Figure 5.8.C and D). These analyses showed the RNA samples were of satisfactory quality to be used as substrates in RT-PCR assays.

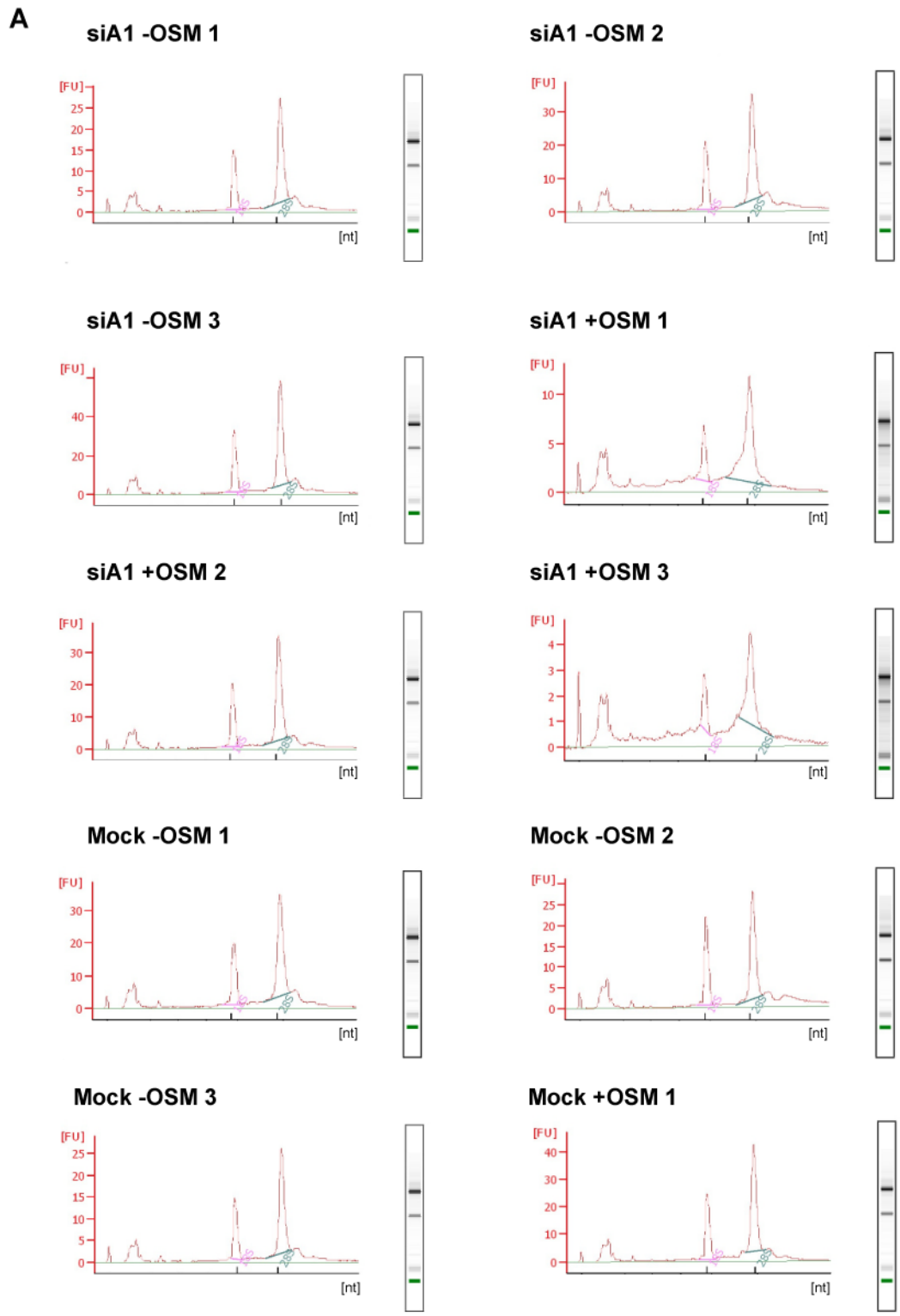


Figure 5.8. Quality analysis of RNA to be used in RT-PCR assays.

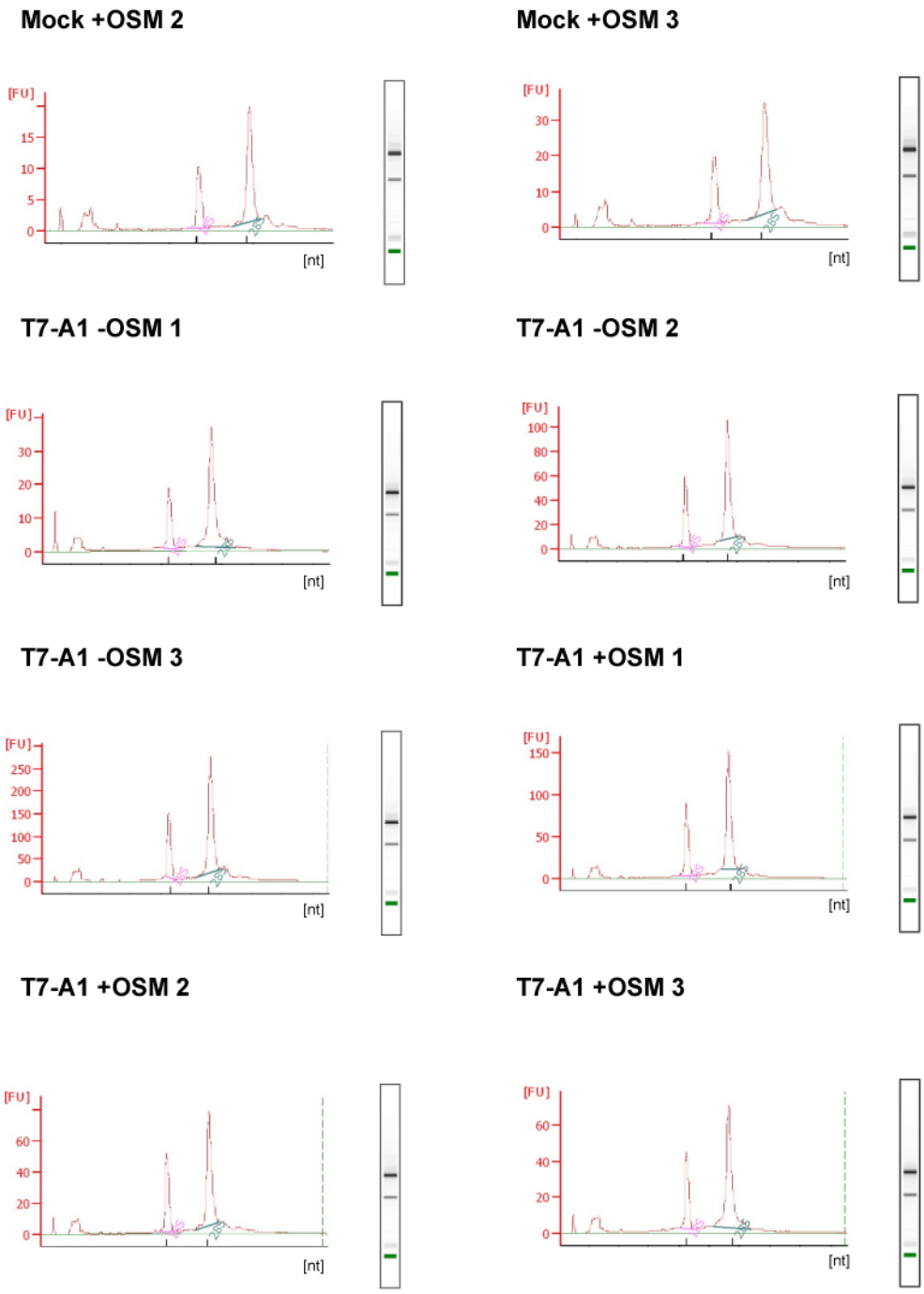


Figure 5.8. Quality analysis of RNA to be used in RT-PCR assays (continued).

B

RNA	Concentration (ng/µl)	28s:18s ratio	RNA integrity number
siA1 –OSM 1	265	1.8	9.7
siA1 –OSM 2	292	1.7	9.7
siA1 –OSM 3	373	1.7	9.8
siA1 +OSM 1	230	3.3	8.3
siAl +OSM 2	228	1.8	9.8
siA1 +OSM 3	172	2.5	7.2
Mock –OSM 1	346	1.7	9.4
Mock –OSM 2	330	1.3	9.8
Mock –OSM 3	278	1.9	9.6
Mock +OSM 1	263	1.6	9.6
Mock +OSM 2	389	1.8	9.6
Mock +OSM 3	350	1.9	9.7
T7-A1 –OSM 1	308	2.7	9.7
T7-A1 –OSM 2	387	1.7	10.0
T7-A1 –OSM 3	391	2.0	9.8
T7-A1 +OSM 1	377	1.7	9.9
T7-A1 +OSM 2	370	1.6	10.0
T7-A1 +OSM 3	322	2.2	9.7

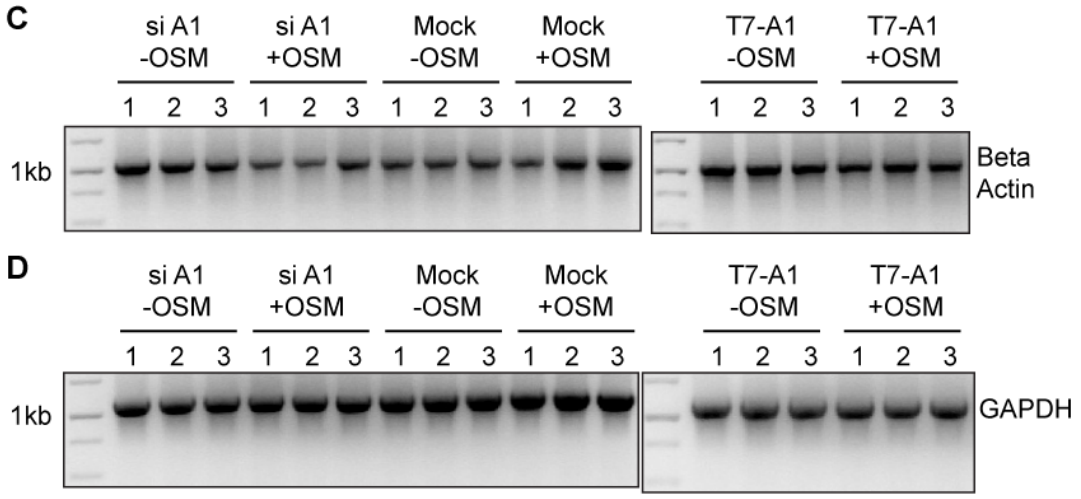


Figure 5.8. Quality analysis of RNA to be used in RT-PCR assays (continued). HeLa cells were transfected by DharmaFECT™ 1 (Dharmacon) with siHnRNPA1 (siA1) or by lipofectamine-2000 (Invitrogen) with T7-hnRNPA1 (T7-A1). Cells were osmotically shocked (OSM) by the addition of 600 mM sorbitol for 2.5 h. Total RNA was prepared from these cells using Trizol (Invitrogen) and analysed by on-chip gel electrophoresis using the Bioanalyzer 2100 (Agilent). The electropherograms, with the 18S and 28S rRNA peaks labelled, and densitometry plots (gel-like images) for each RNA sample are shown in **A**. The concentration, ribosomal RNA ratio and RNA integrity number (RIN) are shown in **B**. The RIN gives an estimate of the integrity of total RNA samples based on the entire electrophoretic trace of the RNA sample and allows RNA samples to be directly compared. 10 indicates a perfect RNA sample without any degradation products, whereas 1 marks a completely degraded sample. To further assess the quality of the RNA, a 1kb fragment of beta-actin (**C**) or GAPDH (**D**) were amplified by RT-PCR (Superscript™ III one-step RT-PCR sytem with Platinum® Taq, Invitrogen) from 500ng of the prepared total RNA samples (Ta 60°C, 40 cycles). Three biological replicates are shown for each condition.

HnRNP A1 prevents intron 3 retention in the CRKL transcript (Figure 5.9.D). Depletion of hnRNP A1 results in detectable levels of a band corresponding to intron retention, although this effect is not consistent across all hnRNP A1 depleted samples which had been osmotically shocked. This may be a result of variations in the extent of hnRNP A1 knockdown. The band corresponding to intron retention is lost upon over-expression of hnRNP A1. In a similar manner, detection of the bands marked with asterisks is inversely correlated to the levels of hnRNP A1; higher intensity bands are detected with depletion of hnRNP A1, and lower intensity bands are detected when hnRNP A1 is over-expressed. These bands were excised, cloned into pGem T-easy and 20 transformants sequenced. Although complete read-throughs of the entire sequence were not obtained, they suggested the presence of one or two additional unannotated exons between exons 3 and 4. In this example inclusion of intron 3, or the additional exons, is clearly dependent upon the levels of hnRNP A1, with the intensity of these bands decreasing as levels of hnRNP A1 increases, and the intensity of the band corresponding to the spliced exon 3-exon 4 product increasing.

HnRNP A1 promotes exons 4 and 5 skipping in the JMJD1C transcript (Figure 5.9.M). Over-expression of hnRNP A1 results in detection of bands corresponding to exon skipping. HnRNP A1 also promotes exon skipping of its own transcripts (Figure 5.9.N, O). This event has been studied previously and shown to be the result of hnRNP A1 binding to an intronic element, CE1, within intron 7, and causing use of the distal 5' ss, therefore promoting exon 8 (or 7B as it is also known) skipping (Chabot *et al.*, 1997). Therefore although I see regulation of this event by hnRNP A1 it is unlikely to be an effect of hnRNP A1 binding to exon 6, but rather to the previously described intronic element.

The transcripts LPHN2 (Figure 5.9.G) and SULT4A (Figure 5.9.I) appear to be differentially regulated upon induction of the stress response. Levels of these transcripts are reduced when cells are stressed, however, over-expression of hnRNP A1 can restore detection of the transcript. Although hnRNP A1 does not appear to regulate the depicted alternative splicing events of these transcripts, these results suggest hnRNP A1 may act to stabilise the transcript in the cytoplasm under times of

environmental stress, as increased levels of hnRNP A1 result in increased amounts of detectable products when compared to mock samples.

HnRNP A1 does not show any detectable modulation of the depicted alternative spliced events in transcripts CDK4 (Figure 5.9.A), RPS16 (Figure 5.9.B), BASP1 (Figure 5.9.C), MIRH1 (Figure 5.9.E), FER (Figure 5.9.J), SLC24A43 (Figure 5.9.L) or FAM38A (Figure 5.9.Q).

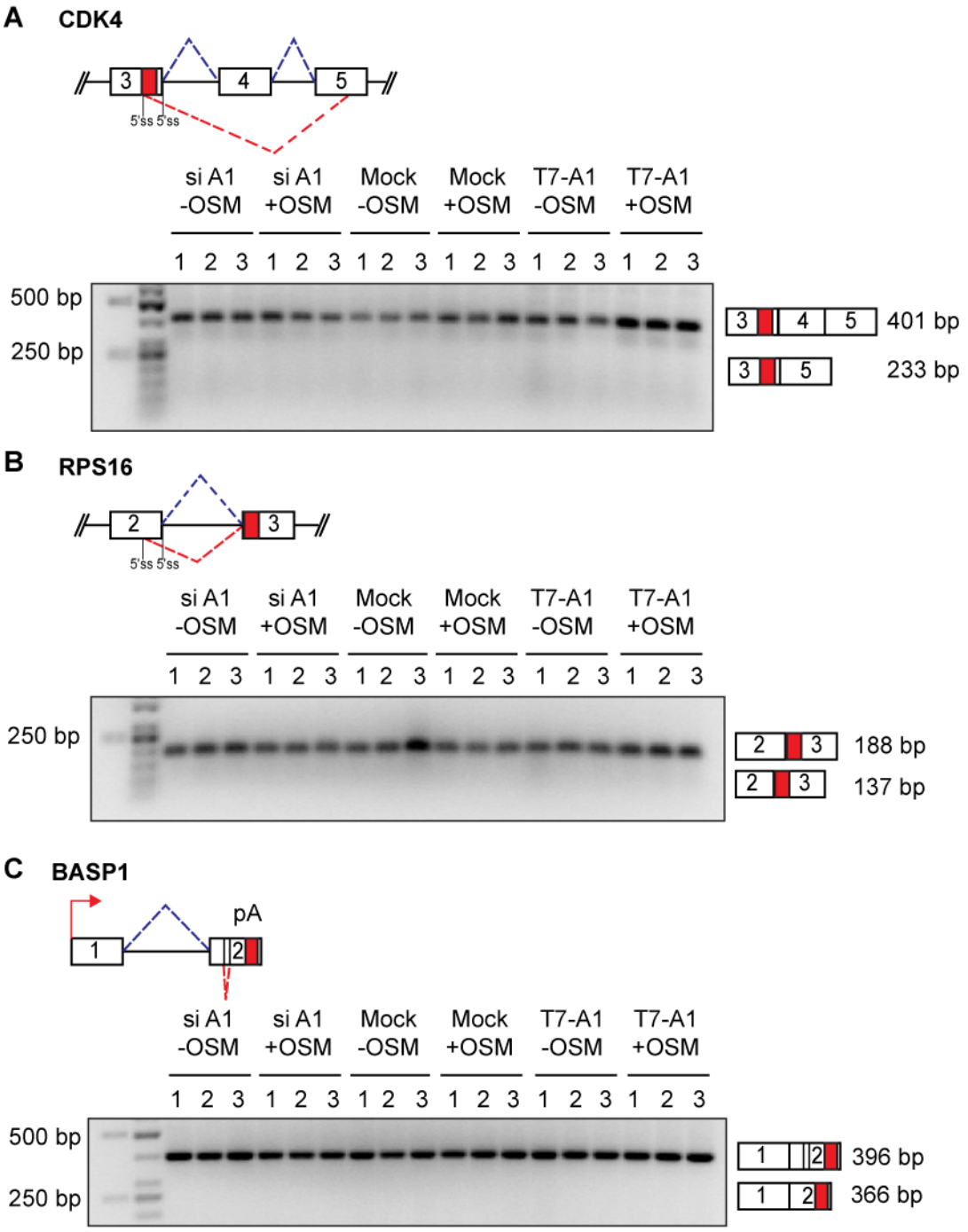


Figure 5.9. Alternative splicing regulation of hnRNP A1 target transcripts. The role of hnRNP A1 in the alternative splicing regulation of sixteen of the transcripts identified by CLIP, where the binding site of hnRNP A1 is located in or near an alternative splicing event, were investigated. Products were amplified by RT-PCR (SuperScript™ III one-step RT-PCR system with Platinum® Taq, Invitrogen) from 500ng of total RNA from hnRNP A1 depleted (siA1), mock, or hnRNP A1 overexpressing (T7-A1) cells, ± osmotic shock (OSM), and analysed on a 1.5% TBE agarose gel stained with ethidium bromide. A schematic of the alternative splicing event is shown above the gel, the red box indicating the location of the CLIP tag. Schematics and sizes of the potential products are shown to the right of the gel. See Materials and Methods for primers used. Three biological replicates are shown for each condition.
* indicates a band of unconfirmed exon composition.

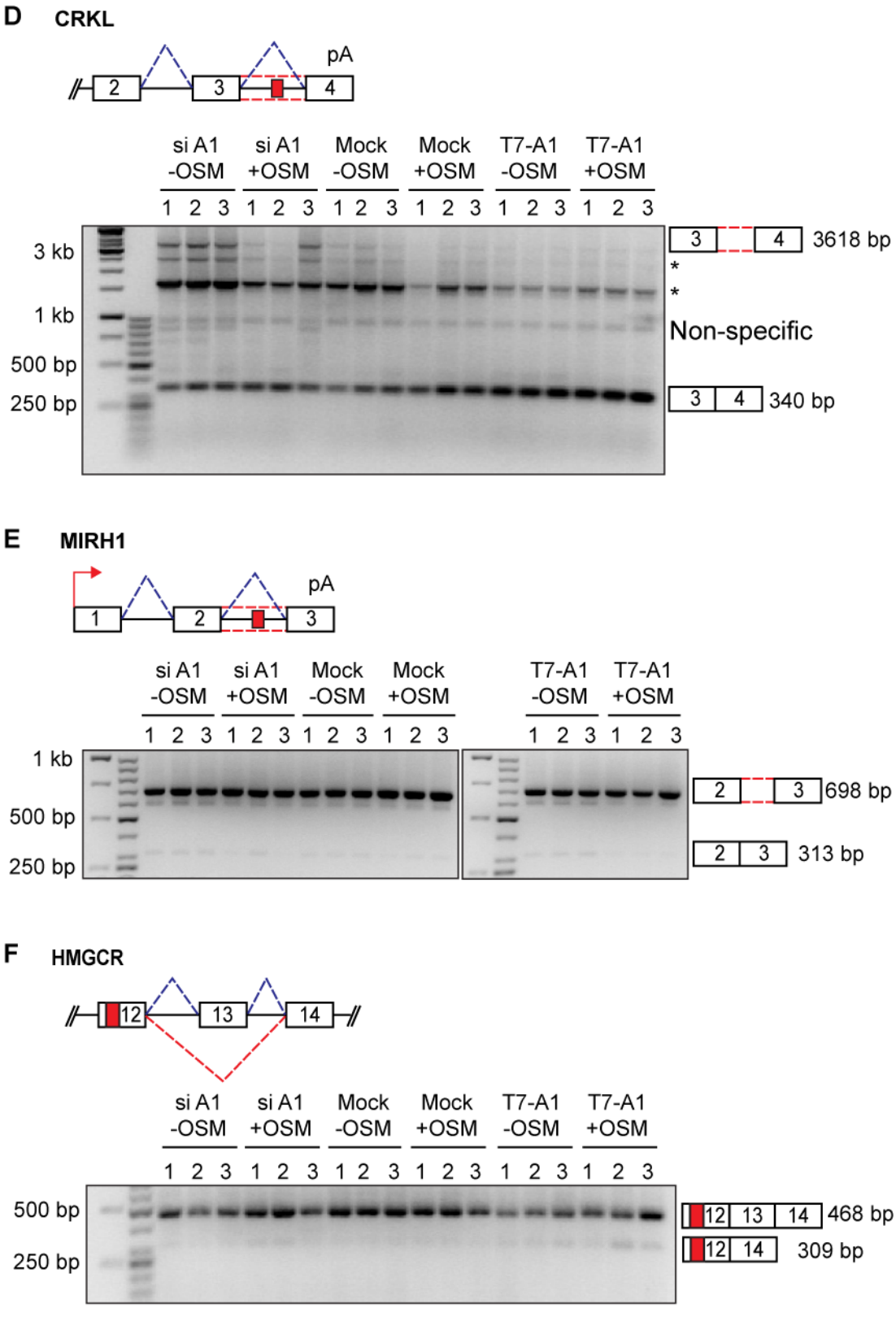


Figure 5.9. Alternative splicing regulation of hnRNP A1 target transcripts (continued).

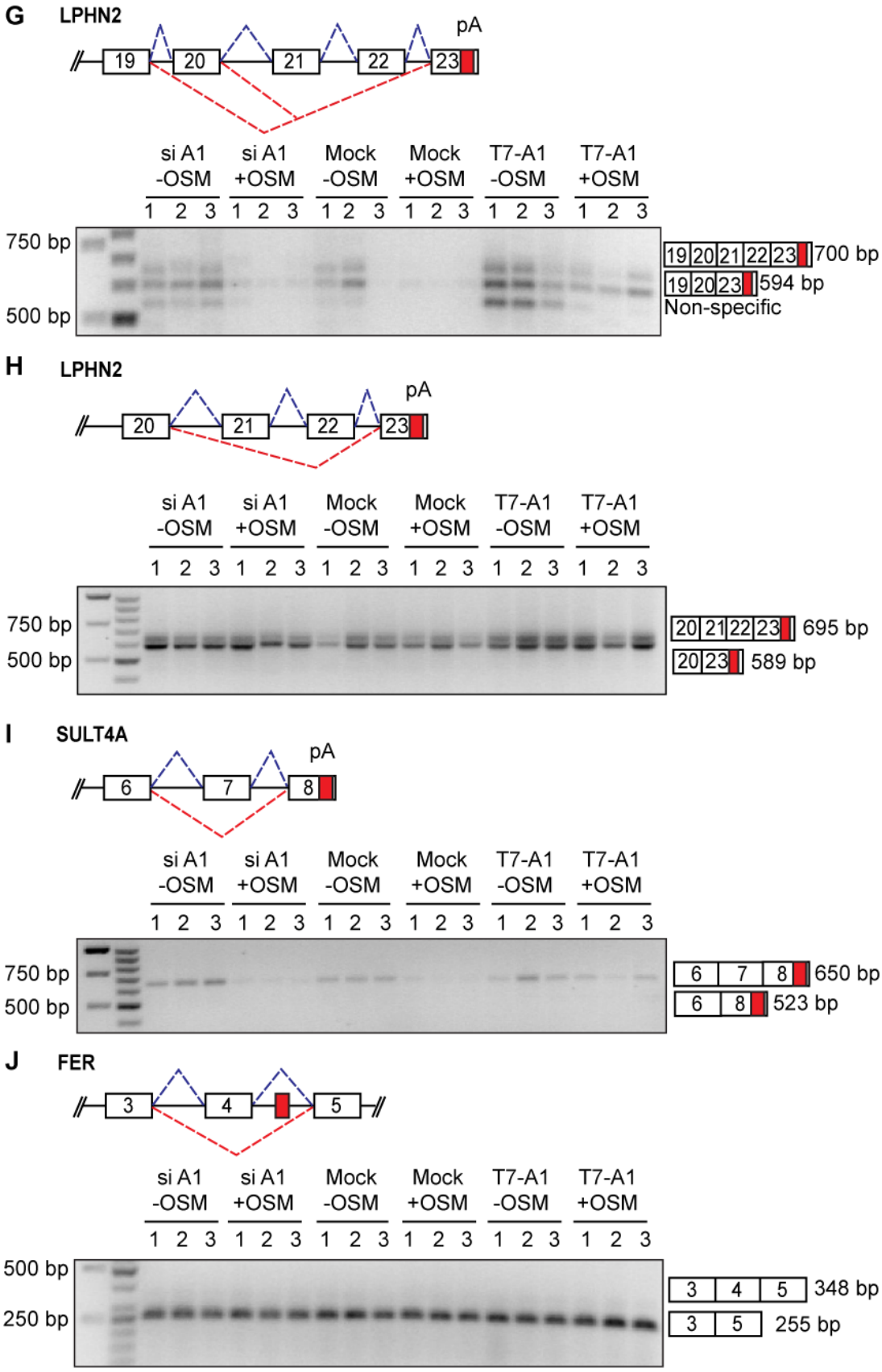


Figure 5.9. Alternative splicing regulation of hnRNP A1 target transcripts (continued).

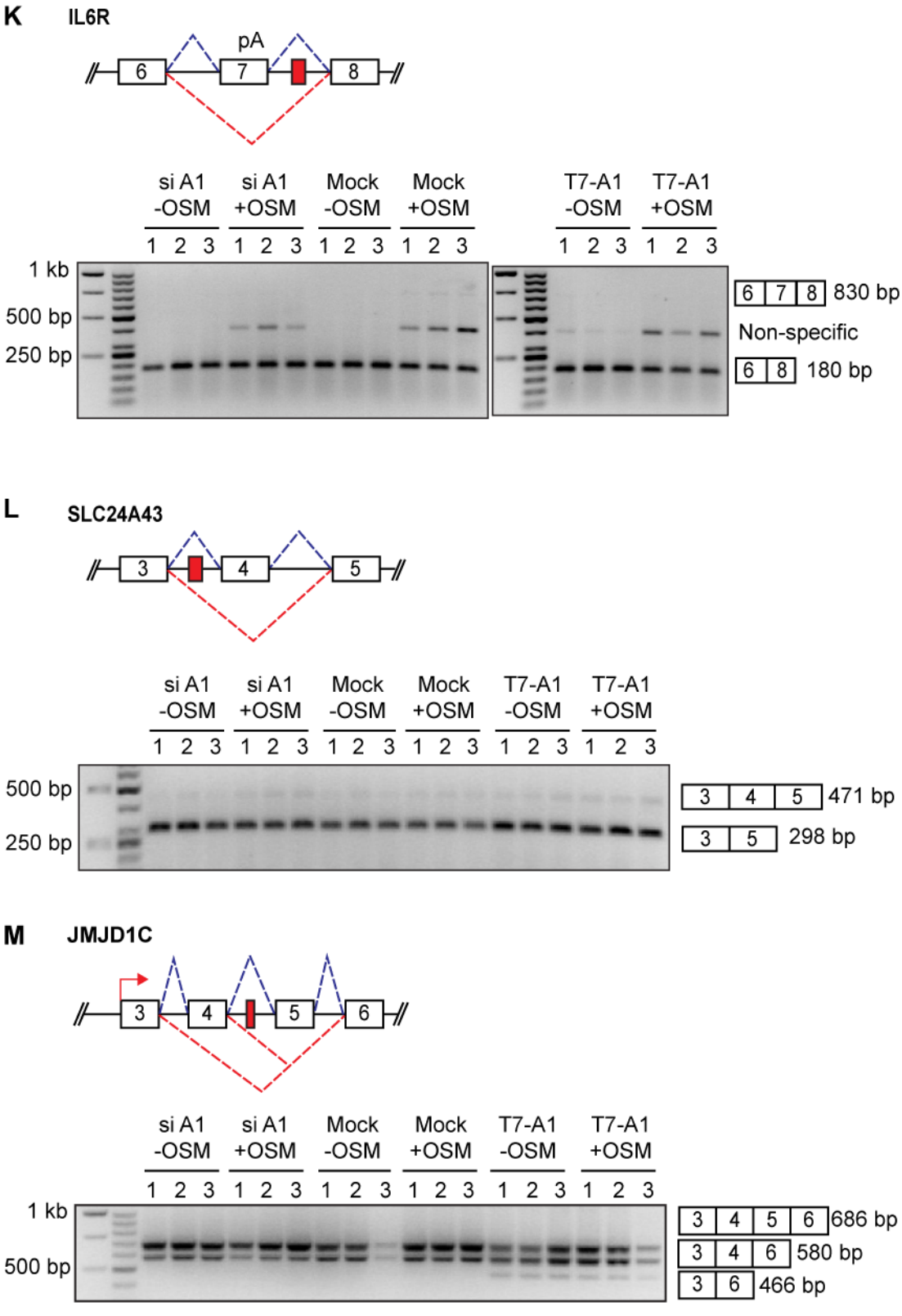


Figure 5.9. Alternative splicing regulation of hnRNP A1 target transcripts (continued).

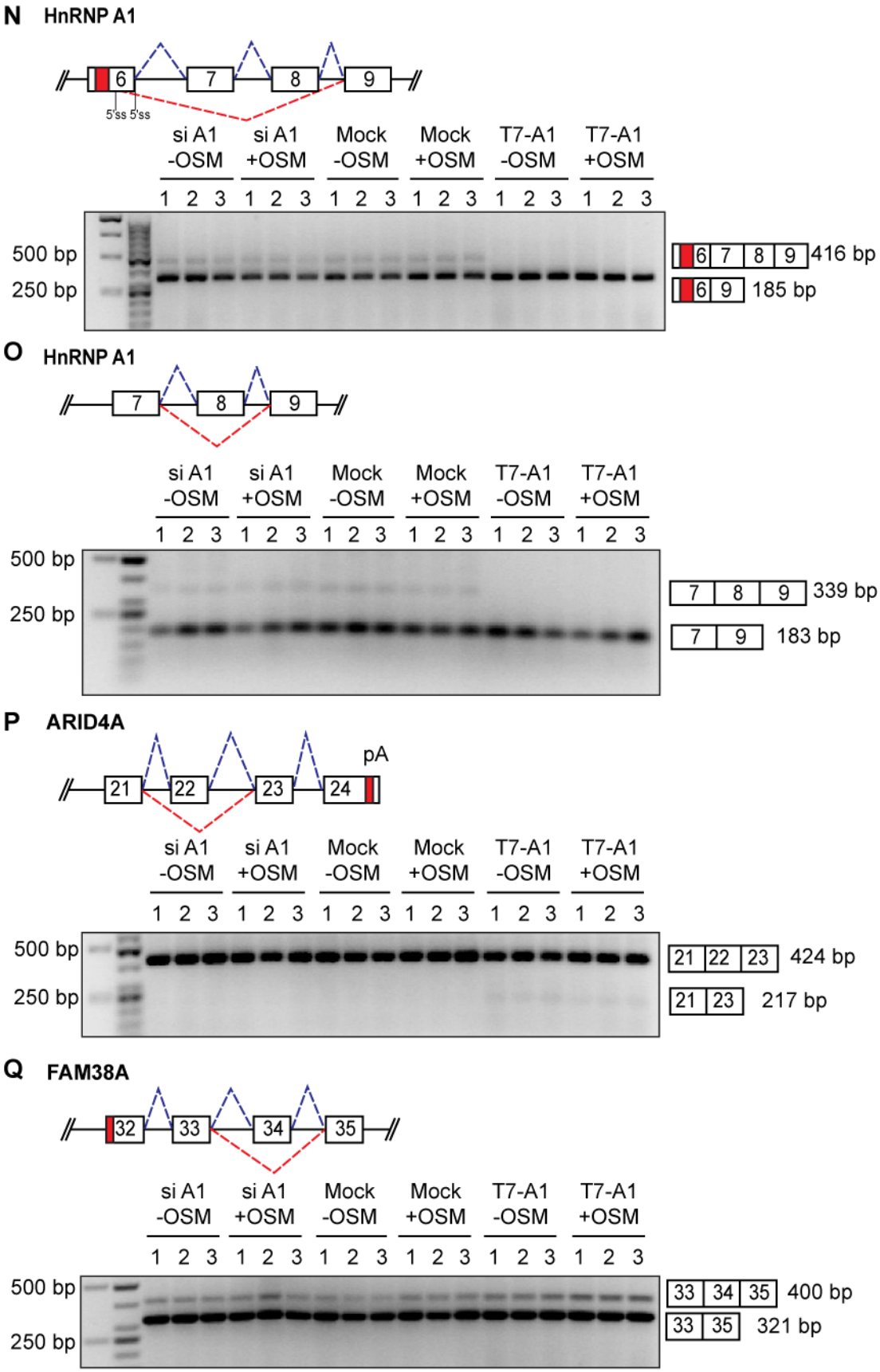
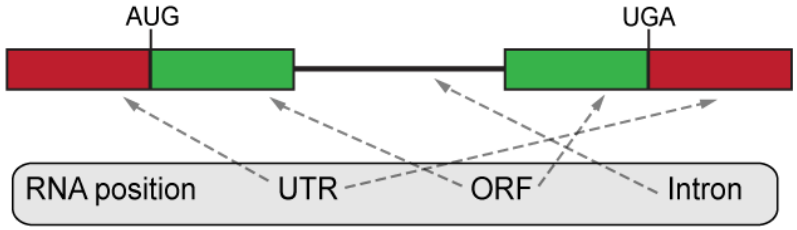


Figure 5.9. Alternative splicing regulation of hnRNP A1 target transcripts (continued).

RNA binding proteins show different binding profiles with respect to gene structure

The protein-coding RNA sequences identified by CLIP as binding to Acinus or hnRNP A1 were analysed in relation to their position within genes. They were classified as binding to 5' or 3' untranslated regions (UTRs), open reading frames (ORFs) and intronic regions. The datasets I generated for Acinus and hnRNP A1 were compared with published data, obtained by CLIP, for Nova (Ule *et al.*, 2003) and SF2/ASF (Sanford *et al.*, 2008) (Figure 5.10.A). Comparisons between proteins suggested they exhibit distinct binding site location preferences, however, in order to determine if these differences were statisitically significant a random dataset was generated. This dataset gives the frequencies of binding to UTRs, ORFs or introns that would be expected if an RNA-binding protein bound randomly to the transcriptome. This was generated by performing an approximation of CLIP *in silico* in collaboration with Colin Semple (MRC Human Genetics Unit). Briefly, the location of randomly generated CLIP-like tags were determined using a transcriptome map constructed from annotated transcripts present in the UCSC database. The random dataset shows that if a protein binds randomly to pre-mRNA transcripts, the predominant binding location is intronic (79%), followed by ORFs (13%) and then UTRs (8%). Comparison of the individual binding site profiles of hnRNP A1, Acinus, SF2/ASF and Nova to the random dataset by Chi-square analyses shows each RNA-binding protein exhibits a highly significantly different binding profile from the random dataset, with all p-values less than 0.0001. Chi-square analyses also show each RNA-binding protein exhibits a significantly different binding profile to each other, which is shown pictorially in Figure 5.10.B.

A



RNA position	UTR	ORF	Intron	p-value
HnRNP A1	9 **10%**	35 **38%**	48 **52%**	<0.0001
Acinus	14 **30%**	30 **64%**	3 **6%**	<0.0001
SF2/ASF	106 **23%**	226 **50%**	122 **27%**	<0.0001
Nova	58 **21%**	38 **13%**	189 **66%**	<0.0001
Random dataset	8547 **8%**	14120 **13%**	83831 **79%**	

B

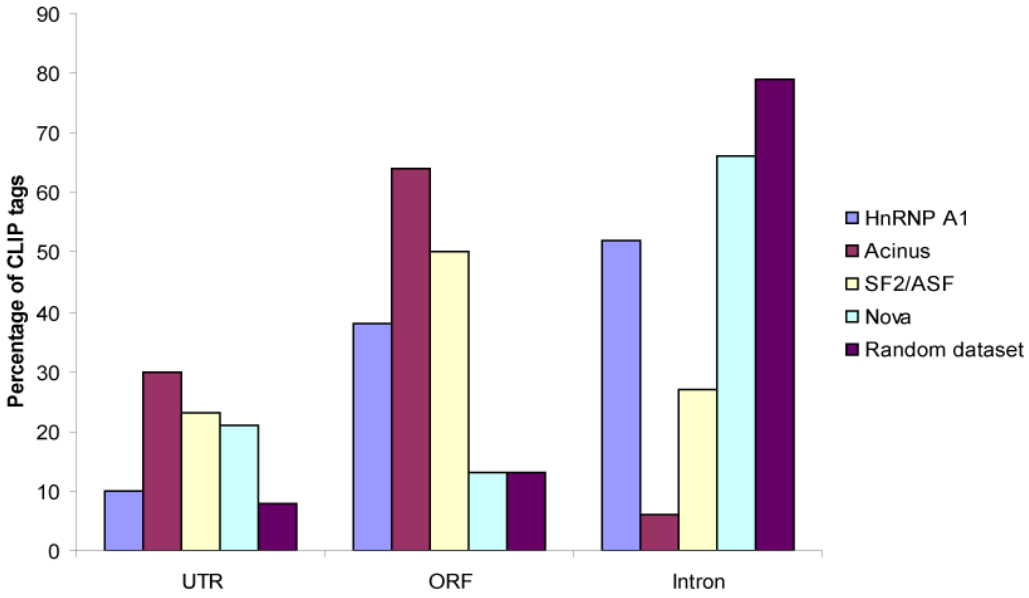


Figure 5.10. RNA binding proteins exhibit distinct binding site location preferences. (**A**) Protein-coding CLIP tags were analysed in relation to their relative positions within genes (5' or 3' untranslated regions (UTR), open reading frame (ORF) or introns). A random dataset was generated by performing an approximation of CLIP *in silico* (see Materials and Methods). This gives the distribution of binding sites, in relation to gene structure, that would be expected if an RNA binding protein bound randomly to the transcriptome. The p-value shown in **A** refers to comparison of the binding site profiles of the individual proteins to the random dataset by Chi squared tests, and shows each protein exhibits a significantly different binding profile. (**B**) The same data displayed in a bar chart, again displaying the differences in binding between proteins. Data for SF2/ASF taken from Sanford *et al.*, 2008, data for Nova taken from Ule *et al.*, 2003. Number of protein coding CLIP tags; HnRNP A1: 92, Acinus: 47; SF2/ASF: 454; Nova: 285. The random dataset was generated by modelling the location of 100 CLIP-like tags and repeating this 1000 times to obtain an average.

This analysis shows the majority (52%) of hnRNP A1 binding sites are intronic, although a substantial amount (38%) are found in ORFs. These binding preferences are in stark contrast to those of Acinus, where the majority (64%) of sites are found in ORFs, with only 6% of binding sites located in introns. SF2/ASF also shows a lower frequency of binding to intronic regions (27%), while Nova binds introns preferentially (66%). These distinct binding profiles may reflect the biological role of the RNA-binding protein, for example hnRNP A1 and Nova are known to bind to ISSs to promote exon skipping, while SF2/ASF has been extensively characterised as binding ESEs and promoting exon inclusion.

5.3. Discussion

This chapter describes the identification of 86 transcripts bound by hnRNP A1, by use of the CLIP technique. CLIP was performed with nuclear and cytoplasmic cellular fractions in the presence and absence of OSM (see Appendix 2), however, due to the small sizes of the datasets, these are presented together here. There were a number of intronic binding sites detected for hnRNP A1, but a large number of exonic sites were also identified (Tables 5.1-5.7). A consensus binding motif of significant e-value could not be detected within the intronic or exonic binding sequences. However, consensus motifs were identified within the nuclear or cytoplasmic sequences using MEME (Figure 5.5). This suggests that hnRNP A1 may exhibit different binding specificities depending upon its cellular localisation, although a larger sample size is required to be confident of this conclusion as the motif found within the cytoplasmic sequences is much less significant than the one found within the nuclear sequences. The nuclear consensus binding motif identified here, (A/G)AGUGA, is a purine-rich sequence similar to the binding motif identified by SELEX (Burd and Dreyfuss, 1994). Additionally, these sequences resemble the consensus sequences of vertebrate 5' and 3' splice sites. This reflects the biological role of hnRNP A1, a predominantly nuclear protein, in nuclear pre-mRNA splicing. HnRNP A1 also exhibits nucleocytoplasmic shuttling, and additionally upon cellular stresses its nuclear localisation is disrupted resulting in cytoplasmic accumulation. The identification of a cytoplasmic consensus binding motif, CTTCCA, which differs substantially from the nuclear consensus binding motif may reflect the role of

hnRNP A1 in cytoplasmic events. These may include regulation of the stability of mRNAs, or trafficking of mRNAs to SGs. Therefore hnRNP A1 may bind different subsets of RNAs depending upon its localisation. How this switch in RNA binding specificity occurs remains to be identified, but it could be modulated by phosphorylation of hnRNP A1 known to occur by the Mnk kinases upon activation of the stress-induced signalling cascade.

Of the sequences identified as bound by hnRNP A1, 34% are located in, or adjacent to an alternative splicing event (Figure 5.3). Several of these alternative splicing events alter the primary structure of the encoded polypeptide (Table 5.7). The resulting protein isoforms may have different biological functions, for example in the IL6R (Interleukin-6 Receptor) transcript exon 7 skipping results in production of the canonical single-pass type I membrane protein, while if exon 7 is included a stop codon is introduced resulting in a shorter, soluble protein lacking the transmembrane domain (Horiuchi *et al.*, 1994). HnRNP A1 also regulates the alternative splicing of its own transcript, which results in the production of protein isoforms exhibiting different activities in splicing reactions. Inclusion of exon 8 (also known as exon 7B) results in the production of the hnRNP A1B isoform, which is less efficient than hnRNP A1 at activating distal 5' splice sites (Mayeda *et al.*, 1994). In addition, exons 7 and 8 skipping results in a protein isoform which lacks the RGG-box motif, known to contribute to RNA binding (Mayeda *et al.*, 1994) (Figure 1.7). HMGCR encodes an enzyme involved in cholesterol synthesis. Exon 13 skipping in the HMGCR transcript results in a 53 amino acid deletion within the catalytic domain, possibly inhibiting its enzymatic actions. Exons 15-56 skipping of the SPTAN1 transcript, encoding a cytoskeletal component, results in the loss of spectrin repeats. These spectrin repeats allow the self-polymerisation of spectrin-proteins which is central to their role as membrane stabilisers and organisers (Kennedy *et al.*, 1994). Therefore deletion of a large number of these repeats due to the alternative splicing event described may impact on the function of SPTAN1 as a cytoskeletal protein. Distinct protein isoforms are also produced as a consequence of the detailed alternative splicing events in transcripts KIF1B, RPS16, TMEM189/UBE2V1 and UBR2.

HnRNP A1 is a known alternative splicing regulator, and I show here that it regulates alternative splicing events in the CRKL and JMJD1C transcripts (Figure 5.9). In the JMJD1C transcript hnRNP A1 promotes exon skipping which correlates with its characterised role in alternative splicing, whereby binding to intronic sites results in the use of distal 5' splice sites promoting exon skipping. The alternative events that hnRNP A1 regulates in the CRKL and JMJD1C transcripts occur in the UTR and have no effect on the primary structure of the encoded protein. However, the event regulated by hnRNP A1 in the CRKL transcript occurs in the 3' UTR, at a site more than 55 nt downstream of the stop codon. Therefore the spliced isoform is expected to be a substrate for NMD as the 'bona fide' stop codon will be deemed as a PTC due to the presence of a downstream EJC.

UTRs contain motifs capable of regulating many aspects of mRNA function, for example nuclear export, cytoplasmic localisation, translational efficiency and stability (Pesole *et al.*, 2001). Alternative splicing contributes to the observed heterogeneity of UTRs, in addition to the use of alternate promoters and polyadenylation sites. Alternative untranslated regions have been demonstrated to determine tissue-specific protein function (Hughes, 2006). Therefore, although these alternative splicing events do not alter the primary structure of the protein, they may alter key motifs within the UTR, which in turn may regulate the expression and biological function of the protein.

I have also shown that hnRNP A1 regulates alternative splicing events within its own transcript (Figure 5.9.N,O). The mechanism of exon 8 (7B) skipping has previously been determined, and is the result of hnRNP A1 binding to an intronic element in intron 7, causing use of the distal 5'ss and promoting exon skipping (Chabot *et al.*, 1997). However, binding of hnRNP A1 to exon 6, as identified by CLIP, may be important in exons 7-8 skipping. Further characterisation of this event, for example by use of *in vitro* splicing assays, is needed to determine the mechanism of this event.

A number of non-coding RNAs, were also identified using CLIP, that bind hnRNP A1. These include the snRNAs U1A, U6 and U12, which are involved in the

splicing reaction, the snoRNA, SNORD3B-2, 7SK snRNA which is involved in transcriptional regulation, and NEAT2, which has been shown to associate with nuclear speckles (Hutchinson *et al.*, 2007). In addition previous work showed pre-miR-18a is a target of hnRNP A1, and hnRNP A1 is required for its processing into a mature microRNA (Guil and Caceres, 2007).

The snRNAs are fairly abundant RNAs, with U1, U2, U4, U5 and U6 snRNAs present at about 10^5-10^6copies per cell, while U11 and U12 snRNAs are slightly less abundant, present at about 10^3-10^4 copies per cell (Montza and Steitz, 1988). Therefore the identification of these interactions may be a consequence of their high abundance.

Interaction of hnRNP A1 with 7SK snRNA has previously been reported (Van Herreweghe *et al.*, 2007). 7SK snRNA controls the activity of the positive transcription elongation factor b (P-TEFb) (Garriga and Grana, 2004; Barboric and Peterlin, 2006; Peterlin and Price, 2006), by sequestering it into kinase inactive complexes (Nguyen *et al.*, 2001, Yang *et al.*, 2001). Inhibition of cellular transcription during times of stress causes disassembly of the 7SK/P-TEFb inactive complexes, and increases the levels of 7SK snRNPs containing hnRNP A1 and other hnRNP proteins (Van Herreweghe *et al.*, 2007). Formation of 7SK/hnRNP complexes allows an increase in nuclear levels of active P-TEFb, possibly by causing disruption of 7SK/P-TEFb complexes, allowing P-TEFb to stimulate transcription. This suggests a role for hnRNP A1 in regulating transcriptional elongation via remodelling of 7SK snRNPs.

HnRNP A1 redistributes to the cytoplasm and accumulates in SGs when cells are under environmental stress (Figure 5.6), and this has previously been shown to change the alternative splicing patterns of the E1A splicing reporter (van der Houven van Oordt *et al.*, 2000). Changes in splicing profiles could have major effects on the fate of the cell, especially as hnRNP A1 is a key regulator in the alternative splicing of apoptotic factors where isoforms can be pro- or anti-apoptotic dependent upon the splicing pattern adopted (Boise *et al.*, 1993; Wang *et al.*, 1994; Shaham and Horvitz, 1996).

Of the transcripts investigated here, only IL6R (Figure 5.9.K) appeared to show an altered splicing pattern in response to osmotic shock. A band of approximately 500 nt was detected in the presence of OSM, but not in the absence of OSM. However, not only was this event not dependent upon hnRNP A1 regulation, as it occurred in hnRNP A1 depleted cells, mock cells and hnRNP A1 over-expressing cells, but upon analysis of the composition of this band it became evident it was non-specific and hit back to a random genomic location. This is puzzling as the primers used in these assays were designed carefully and those which were predicted to amplify non-specific fragments were discarded. It is also intriguing that this band is only amplified in cells which had been osmotically shocked.

Therefore although none of the events investigated here show a change in splicing pattern (Figure 5.9), two transcripts (LPHN2 and SULT4A) appear to be differentially regulated upon induction of the stress response. Although it is puzzling that a stress dependent response is seen when amplifying LPHN2 from exon 19, but this is lost when amplifying from exon 20 (Figure 5.9.G, H). In both mock and hnRNP A1 depleted cells the transcripts are undetected in the presence of OSM, while the transcripts are detected in the presence of OSM, upon over-expression of hnRNP A1. This suggests a role for hnRNP A1 in regulating the stability of mRNAs during the stress response, and perhaps ensuring their localization to SGs. Indeed, previous studies have indicated that cytoplasmic hnRNP A1 binds reiterated AUUUA sequences that have been shown to modulate mRNA turnover and translation (Hamilton *et al.*, 1997). Therefore hnRNP A1 may function like other RNA stability regulators, such as HuR and TTP known to be present in SGs (Kedersha and Anderson, 2002), to ensure a particular subset of RNAs is sorted efficiently in times of stress.

It is interesting to see the distinct binding profiles of the different RNA-binding proteins when analysed with respect to gene structure (Figure 5.10). The random dataset was generated by modelling CLIP *in silico*, although there are a number of caveats to this procedure which must be noted. The majority of these involve the generation of a transcriptome map which was constructed from annotated transcripts present in the UCSC database in the hg 18 2006 assembly. However, we cannot

accurately represent the transcriptome due to unknown variables, such as non-annotated transcripts, non-coding transcripts (which were eliminated for this exercise), transcript abundance (each transcript was modelled as being equally abundant), plus cell line specific abundancies which were also over-looked. Despite these caveats it gave an approximation of the expected distribution a protein would exhibit if it bound by chance to the protein-coding transcriptome, and as expected this was predominantly intronic (79%) due to the prevalence of intronic regions in human coding genes.

Therefore all of the RNA-binding proteins show binding preferences that are significantly different to what might be expected by chance alone. Surprisingly, hnRNP A1 binds roughly 50% each to intronic and exonic elements. HnRNP A1, along with other hnRNPs, repress splicing by binding silencer elements. Most described splicing silencers are intronic elements, although several ESSs have been reported (reviewed by (Cartegni *et al.*, 2002)). Additionally, putative binding sites for hnRNP A1 were found to be more abundant in introns (1.2 sites per 1000 bp) than exons (0.35 binding sites per 1000 bp) when six genes were examined (Blanchette and Chabot, 1999). This demonstrates that use of a more global analysis of RNA binding protein targets, such as CLIP, can provide more information than studying a few specific interactions.

Chapter 6: Discussion and Future Work

The work presented here has used the CLIP technique to identify RNA targets directly bound to Acinus and hnRNP A1. This was undertaken to identify endogenous transcripts these proteins bind, in order to increase our understanding of their role in the post-transcriptional regulation of gene expression.

Although hnRNP A1 has been extensively characterised as an RNA processing factor, acting in many stages of RNA metabolism, Acinus was largely uncharacterised. Chapter 3 describes the further characterisation of Acinus, and shows Acinus is a factor involved in both apoptosis and RNA processing. Furthermore, alternative splicing assays of transcripts identified as targets of Acinus by CLIP, shows modulation of the levels of Acinus affects the alternative splicing of two transcripts (MYST1 and CACHD1, Figure 4.6).

Acinus was originally identified as a factor which caused apoptotic chromatin condensation following caspase-3 cleavage (Sahara *et al.*, 1999). The authors of this paper identified a region with homology to the RRM of the Drosophila splicing regulator Sxl (Figures 1.13, 3.6). The original interest in Acinus was based on the presence of the intact RRM within the caspase-cleaved product (Acinus (987-1093), Figure 1.12), and led to the hypothesis that RNA binding may be necessary for Acinus' function in the apoptotic pathway.

Despite indications that Acinus was an RNA processing factor; it is a component of the spliceosome (Rappsilber *et al.*, 2002; Zhou *et al.*, 2002), the EJC (Tange *et al.*, 2005) and part of the ASAP complex (Schwerk *et al.*, 2003), this was unconfirmed. Here I show that Acinus directly binds polyadenylated RNA (Figure 3.6), is localised to nuclear speckles (Figure 3.3) and regulates the alternative splicing of the MYST1 and CACHD1 transcripts (Figure 4.6). This data provides convincing evidence that Acinus is a splicing factor.

I also demonstrate that over-expression of Acinus S does induce apoptotic chromatin condensation in a concentration-dependent manner (Figure 3.11). The manner by which transient expression of T7-Acinus S causes apoptosis in unclear. Previous studies in cell-free systems have shown that, unlike the caspase-cleaved Acinus (987-1093) protein, the full-length Acinus isoforms do not induce chromatin

condensation (Sahara *et al.*, 1999). It is possible that due to the complexities of the regulatory networks controlling apoptosis *in vivo*, changes in levels of key factors may push the cell towards apoptosis. If this is the case then the levels of Acinus must be carefully regulated as I have shown it is ubiquitously expressed (Figure 3.5). Alternatively, increased levels of Acinus S may increase the probability of interaction with residual caspases, and result in production of low levels of the cleaved Acinus protein which can induce apoptosis. Although induction of apoptosis by over-expression of Acinus S is statistically significant when compared with control cells, it occurs at substantially lower levels than that observed when apoptosis is induced with STS (Figure 3.11). This may suggest that Acinus acts in conjunction with other factors to mediate apoptosis. In support of this hypothesis depletion of Acinus does not affect the potential of a cell to undergo apoptosis (Figure 3.11), indicating Acinus has redundant functions in the apoptotic pathway, which can also be mediated by factors such as CAD and AIF.

This highlights the difficulties of dissecting the apoptotic process. It may be necessary to use cell-free systems to analyse the function of Acinus in apoptosis. This has the advantage of simplifying the complexities of cellular systems and allows identification of factors responsible for specific events. Cell-free systems have been widely used in the dissection of the central mechansims of apoptosis, and were employed in the original identification of Acinus (Sahara *et al.*, 1999).

To determine if the RNA binding ability is critical in mediating chromatin condensation, Acinus (987-1093) was cloned into the pCGT7 vector. Additionally, essential amino acids in the RNP motifs within the RRM, based upon previously described mutations which disrupted RNA binding (Caceres and Krainer, 1993), were mutagenised in order to generate RNA binding deficient mutants. By comparing the number of cells which showed morphology indicative of apoptosis upon transient transfection of wild-type Acinus (987-1093) versus the RNA binding deficient mutants, it was hoped to determine if RNA binding may play a role in inducing chromatin condensation. However, I encountered problems both expressing and visualising these T7-Acinus (987-1093) proteins. Initial results indicated the proteins were expressed when detected by Western blot, but expression levels

between the mutants varied, and I could not achieve convincing data to demonstrate they did not bind RNA. Immunostaining also proved problematic, although the observation of large amounts of cellular debris upon transient transfection of T7-Acinus (987-1093) which is not seen in other cell preparations, could indicate the cells have already undergone apoptosis. This debris could represent apoptotic bodies containing cellular material, which are engulfed by other cells and digested via the lysosomal pathway. If this is the case, it suggests that over-expression of Acinus (987-1093) causes rapid induction of apoptosis.

Generation of the same mutations in the pCGT7-Acinus S construct, and cell-counting assays may determine if RNA binding is necessary for Acinus' role in apoptosis. However, it is also important to determine if the RRM is responsible for RNA binding. The RS domain may also mediate this interaction, as previously reported (Nikolakaki *et al.*, 2008) and therefore the cleaved Acinus protein (987-1093) may actually be deficient in RNA binding.

It would also be interesting to generate a caspase-resistant mutant, by site-directed mutagenesis of the DEXD caspase-3 cleavage site in the pCGT7-Acinus S construct, and determine if this construct still causes apoptosis in a concentration-dependent manner. If it does it would suggest that Acinus does not require cleavage to be active in apoptosis, but over-expression is sufficient. Therefore regulation of Acinus protein expression must be highly important in preventing cell death, as it is such an ubiqutous protein (Figure 3.5).

Several of the transcripts identified as binding Acinus by CLIP, encode products involved in apoptosis (ACIN1, F2R, MZF1 and MYST1). These interactions may be important in the regulation of apoptosis. For example, over-expression of MZF1 inhibits apoptosis (Hromas *et al.*, 1996, Robertson *et al.*, 1998). Acinus may regulate the levels of MZF1 by binding its transcript, and thereby exert control over the apoptotic pathway. MYST1 is a histone acetyltransferase, which also acetylates p53, a central factor of the apoptotic pathway, and this modification causes the expression of apoptotic genes (Sykes *et al.*, 2006). Futhermore, MYST1 mutant embryos exhibit abnormal chromatin morphology before undergoing apoptosis (Thomas *et al.*, 2008). Therefore MYST1 may regulate chromatin architecture, and loss of function may

contribute to the chromatin condensation observed in apoptosis. I show Acinus can modulate the alternative splicing profile of the MYST1 transcript (Figure 4.6). Acinus promotes intron 11 retention which results in the production of a protein isoform with a different C-terminus to the canonical isoform. There are no reports of the function of this alternatively spliced isoform, but it would be interesting to determine if this isoform can rescue the null phenotype observed. If it cannot then it is possible that regulation of MYST1 isoforms by Acinus may affect the apoptotic process.

Regulation of the cell cycle is also an important aspect in both cell proliferation and apoptosis. This is required to acheive homeostasis in multicellular organisms, as well as in instances that require apoptotic degradation of cells or tissues. Several of the identified targets of Acinus encode products involved in the cell cycle (ANAPC1, F2R and GNL3). This extends the possible regulatory network that Acinus acts upon in order to regulate the induction of apoptosis. However, these interactions were detected using T7-Acinus S and may not occur with the cleaved Acinus (987-1093) protein. The function of Acinus as a regulator of these transcripts is likely to be separate from its role in inducing chromatin condensation once apoptosis has been induced. This is likely as apoptosis generally proceeds rapidly and switching alternative splice profiles is unlikely to be an effective mode of regulation, however, it may play a more important role in slower forms of cell death such as neurodegeneration.

Therefore I suggest a model whereby Acinus regulates the expression of factors, involved in the cell cycle, cell proliferation and apoptosis, by binding their transcripts and either modulating their alternative splicing or stability. In addition Acinus functions during the apoptotic process, as a direct target of caspase-3, and mediates chromatin condensation together with other downstream factors of caspase-3. Although the mechanism by which this is acheived is unknown, I hypothesise that Acinus is a chromatin associated protein which places it in an ideal location to facilitate interactions causing chromatin condensation.

HnRNP A1 is involved in the stress response, as previously discussed, and upon cellular stresses it is redistributed to the cytoplasm where it accumulates in stress

granules. Here I show that although hnRNP A1 does not appear to affect the splicing profile of any of the transcripts identified by CLIP during the stress response, it does appear to stabilise the LPHN2 and SULT4A transcripts (Figure 5.9). Previous studies have shown that the cytoplasmic RNA binding ability of hnRNP A1 is important during the stress response (Guil *et al.*, 2006), which may involve promotion of mRNA stability, or involvement in RNA localisation to stress granules.

In response to stress eukaryotic cells shut down protein synthesis to conserve energy for the repair of stress-induced damage. Stalled preinitiation complexes and their associated mRNAs are dynamically routed to SGs. This process is known to involve the RNA-binding proteins TIA-1 and TIAR. (Kedersha *et al.*, 2000; Kedersha *et al.*, 2002). In order to determine if hnRNP A1 is also required to transport RNAs to SGs I proposed to study the localisation of targets identified by CLIP and determine if this changed upon the induction of stress, and whether this was dependent upon hnRNP A1. The localisation of RNAs can be determined by RNA FISH (fluorescent *in situ* hybridisation), and preliminary analyses showed two transcripts (HN1 and TMEM189/UBE2V1) did localise to SGs upon treatment with OSM. By depleting cells of hnRNP A1 by RNAi, and performing the same analysis, it may be possible to determine if hnRNP A1 is necessary to route these transcripts to SGs.

A large portion of this thesis has concentrated on use of the CLIP technique to identify *in vivo* targets of Acinus and hnRNP A1. This has allowed the identification of novel targets of these RNA-binding proteins, and the presence of binding sites in or near sites of alternative splicing (Figures 4.3, 5.2) suggested Acinus and hnRNP A1 may regulate the splicing profile of these transcripts. Indeed, analysis of these alternative events showed both Acinus and hnRNP A1 could modulate the alternative splicing of specific transcripts (Figures 4.6, 5.9).

Some of the alternative splicing events in the transcripts identified by CLIP result in changes in the primary structure of the encoded protein (Tables 4.8, 5.7). However, a number of these events result in the production of PTCs, therefore marking these transcripts as substrates for the NMD pathway. NMD is an RNA surveillance pathway whereby errors in gene expression are degraded before these

transcripts are translated into potentially toxic truncated proteins (Chang *et al.*, 2007). Estimates are 35% of alternative spliced isoforms are subject to NMD (Lewis *et al.*, 2003; Hillman *et al.*, 2004; Baek and Green, 2005), and studies of these events has led to the hypothesis that these events may exploit NMD to achieve quantitative post-transcriptional regulation (AS-NMD) (McGlincy and Smith, 2008). All members of the SR protein family undergo unproductive splicing, or AS-NMD, as a form of regulation (Lareau *et al.*, 2007). Therefore it is possible that these alternative splicing events which introduce PTCs, occurring in transcripts bound by both Acinus and hnRNP A1, may represent important regulatory events.

Performing CLIP with two RNA-binding proteins allows for comparison of their binding sites and transcripts, which may reflect the differences in their biological function. There was no overlap between the subsets of RNAs identified as targets of Acinus and hnRNP A1, although larger sample sizes are required before any conclusions can be made. Additionally, the binding locations of the two proteins, with respect to gene structure were significantly different (Figure 5.10).

Acinus binds almost exclusively to exonic regions, while hnRNP A1 binds with approximately equal frequencies to exons and introns. This is surprising as hnRNP A1 binding sites have been reported to be more abundant in introns than exons (Blanchette and Chabot, 1999). This data may suggest hnRNP A1 binds more extensively to exonic regions to regulate alternative splicing than previously thought, or these interactions mediate other aspects of RNA metabolism that hnRNP A1 is known to function in, such as nucleocytoplasmic export, localisation and stability.

Acinus also exhibits a high frequency of binding to UTRs, which is not seen by hnRNP A1, when compared to the random dataset (Figure 5.10). Binding to UTRs may allow the regulation of a number of aspects of mRNA function as discussed earlier, for example stability or translational efficiency, and this may indicate a potential mechanism by which Acinus may regulate RNA metabolism.

These comparisons between Acinus and hnRNP A1 illustrate some of the advantages of the CLIP technique. Methods like CLIP have the potential to aid understanding of the specificity and functions of RNA-binding proteins by making

connections between trans-acting RNA binding proteins and the biological processes they regulate (Ule *et al.*, 2003). CLIP also aids the identification of cis-acting RNA elements, as positional information is acquired of the protein-RNA interaction. In addition CLIP has the potential of identifying non-coding RNA targets of RNA binding proteins. Work from our laboratory identified miR18a as an RNA target of hnRNP A1 (Guil and Caceres, 2007) as a result of a CLIP screen. In the light of recent data from the ENCODE project where, it is estimated, 93% of genomic sequences are capable of being transcribed (The Encode Project Consortium, 2007) it is important to determine what auxiliary factors bind these ncRNAs, and as a result elucidate their function.

However, despite these advantages of CLIP there are a number of associated problems. The technique itself is technically difficult, involving a number of steps, and optimisation is required for each protein studied. CLIP is most suitable to be used on small to medium sized proteins, such as hnRNP A1 or SF2/ASF, as resolution of the protein-RNA complexes by SDS-PAGE can prove difficult with larger proteins. This was the case with Acinus; T7-Acinus S migrates at about 100 kDa, and even using low percentage acrylamide gels it could be difficult to detect a shift of 20 kDa upon RNase digestion.

CLIP is also very inefficient, the amount of useable sequence data compared to the number of transformants sequenced is very low. This is due to both empty inserts, contaminating sequences and an abundance of ribosomal RNA (rRNA) sequences. In these analyses rRNAs were discarded, as due to the high abundance of rRNAs in the cell these may not be real targets. This also highlights the need for validation of the transcripts identified by CLIP. A few transcripts were validated here by means of alternative splicing assays, however, the remainder have not been validated and need to be by some independent method, such as IP-RT-PCR or gel-shift analyses, to determine these are ‘bona fide’ targets. These techniques also have the advantage of further characterising the precise site of protein-RNA interaction.

In addition, CLIP demands the use of bioinformatic analyses of the sequences generated, as manual annotation is labourious and time-consuming.

Therefore, although CLIP offers several advantages in identifying regulatory networks of interactions of RNA-binding proteins, it is also a difficult and lengthy screen which requires further validation and characterisation of the identified transcripts. However, advancements in bioinformatics and high-throughput sequencing platforms has led to a modification of the CLIP technique described as HITS (HIgh Throughput Sequencing)-CLIP, whereby millions of RNAs can be sequenced and their genomic position characterised. This has the advantage of requiring less experiments, providing a large amount of data, and also allows some kind of confidence that the binding sites are 'bona fide' by the frequency they are obtained. However, this requires even more bioinformatic analysis of the data and effectively generates a catalogue of targets which would require further characterisation to ascertain the true biological function, but none-the-less it provides a valuable resource which researchers can use to direct their studies.

References

Adams, J. M. (2003) Ways of dying: multiple pathways to apoptosis. Genes Dev. **17**, 2481-2495

Allemand, E., Guil, S., Myers, M., Moscat, J., Cáceres, J.F. and Krainer, A.R. (2005) Regulation of heterogenous nuclear ribonucleoprotein A1 transport by phosphorylation in cells stressed by osmotic shock. Proc Natl Acad Sci U S A. **102**, 3605-3610

Amrein, H., Gorman, M. and Nothiger, R. (1988) The sex-determining gene tra-2 of Drosophila encodes a putative RNA binding protein. Cell **55,** 1025-1035

Baek, D. and Green, P. (2005) Sequence conservation, relative isoform frequencies, and nonsense-mediated decay in evolutionarily conserved alternative splicing. Proc Natl Acad Sci U S A. **102**, 12813-12818

Bailey, T.L. and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol. **2**, 28-36

Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Seraphin, B. and Le Hir, H. (2005) The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. Nat Struct Mol Biol. **12**, 861-869

Barnard, D.C. and Patton, J.G. (2000) Identification and characterization of a novel serine-arginine-rich splicing regulatory protein. Mol Cell Biol. **20**, 3049-3057

Barnard, D.C., Li, J., Peng, R. and Patton, J.G. (2002) Regulation of alternative splicing by SRrp86 through coactivation and repression of specific SR proteins. RNA. **8**, 526-533

Bauren, G. and Wieslander, L. (1994) Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. Cell **76**, 183-192

Bedard, K.M., Daijogo, S. and Semler, B.L. (2007) A nucleo-cytoplasmic SR protein functions in viral IRES-mediated translation initiation. EMBO J. **26**, 459-467

Beil, B., Screaton, G. and Stamm, S. (1997) Molecular cloning of htra2-beta-1 and htra2-beta-2, two human homologs of tra-2 generated by alternative splicing. DNA Cell Biol. **16,** 679-690

Berget, S.M., Moore, C. and Sharp, P.A. (1977) Spliced segments at the 5' terminus of adenovirus 2 late mRNA. Proc.Natl.Acad.Sci.U.S.A **74,** 3171-3175

Bessonov, S., Anokhina, M., Will, C.L., Urlaub, H. and Luhrmann, R. (2008) Isolation of an active step I spliceosome and composition of its RNP core. Nature **452,** 846-850

Beyer, A.L. and Osheim, Y.N. (1988) Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. Genes Dev. **2,** 754-765

Birmingham, A., Anderson, E., Sullivan, K., Reynolds, A., Boese, Q., Leake, D., Karpilow, J. and Khvorova, A. (2007) A protocol for designing siRNAs with high functionality and specificity. Nature Protocols **2**, 2068-2078

Birney, E., Kumar, S. and Krainer, A.R. (1993) Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. Nucleic Acids Res. **21,** 5803-5816

Black, D.L. (2003) Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem. **72**, 291-336

Blanchette, M. and Chabot, B. (1999) Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. EMBO J. **18**, 1939-1952

Blanchette, M., Green, R.E., Brenner, S.E. and Rio, D.C. (2005) Global analysis of positive and negative pre-mRNA splicing regulators in Drosophila. Genes Dev. **19,** 1306-1314

Blaustein, M., Pelisch, F., Coso, O.A., Bissell, M.J., Kornblihtt, A.R. and Srebrow, A. (2004) Mammary epithelial-mesenchymal interaction regulates fibronectin alternative splicing via phosphatidylinositol 3-kinase. J.Biol.Chem. **279,** 21029-21037

Blaustein, M., Pelisch, F., Tanos, T., Munoz, M.J., Wengier, D., Quadrana, L., Sanford, J.R., Muschietti, J.P., Kornblihtt, A.R., Caceres, J.F., Coso, O.A. and Srebrow, A. (2005) Concerted regulation of nuclear and cytoplasmic activities of SR proteins by AKT. Nat.Struct.Mol.Biol. **12,** 1037-1044

Blencowe, B.J., Bowman, J.A., McCracken, S. and Rosonina, E. (1999) SR-related proteins and the processing of messenger RNA precursors. Biochem.Cell Biol. **77,** 277-291

Blencowe, B.J., Issner, R., Nickerson, J.A. and Sharp, P.A. (1998) A coactivator of pre-mRNA splicing. Genes Dev. **12,** 996-1009

Boatright, K. M. and Salvesen, G. S. (2003) Mechanisms of caspase activation. Curr. Opin. Cell Biol., **15**, 725-731.

Boggs, R.T., Gregor, P., Idriss, S., Belote, J.M. and McKeown, M. (1987) Regulation of sexual differentiation in D. melanogaster via alternative splicing of RNA from the transformer gene. Cell **50,** 739-747

Boise, L.H., González-García, M., Postema, C.E., Ding, L., Lindsten, T., Turka, L.A., Mao, X., Nuñez, G. and Thompson, C.B. (1993) bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. Cell, **74**, 597-608.

Boucher, L., Ouzounis, C.A., Enright, A.J. and Blencowe, B.J. (2001) A genome-wide survey of RS domain proteins. RNA.**7,** 1693-1701

Bourgeois, C.F., Lejeune, F. and Stevenin, J. (2004) Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. Prog.Nucleic Acid Res.Mol.Biol. **78,** 37-88

Buratti, E., Muro, A.F., Giombi, M., Gherbassi, D., Iaconcig, A. and Baralle, F.E. (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,** 1387-1400

Burd, C.G. and Dreyfuss, G.(1994) RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. EMBO J. **13**, 1197-1204

Caceres, J. F. and Krainer, A. R. (1993) Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. EMBO J. **12**, 4715-4726

Caceres, J.F., Misteli, T., Screaton, G.R., Spector, D.L. and Krainer, A.R. (1997) Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. J.Cell Biol. **138,** 225-238

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

Caceres, J.F., Stamm, S., Helfman, D.M. and Krainer, A.R. (1994) Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. Science **265,** 1706-1709

Cao, W., Jamison, S.F. and Garcia-Blanco, M.A. (1997) Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. RNA **3,** 1456-1467

Cartegni, L. and Krainer, A.R. (2002) Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. Nat.Genet. **30,** 377-384

Cartegni, L., Chew, S.L. and Krainer, A.R. (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat.Rev.Genet. **3,** 285-298

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

Carter, M.S., Li, S. and Wilkinson, M.F. (1996) A splicing-dependent regulatory mechanism that detects translation signals. EMBO J. **15**, 5965-5975

Cavaloc, Y., Bourgeois, C.F., Kister, L. and Stevenin, J. (1999) The splicing factors 9G8 and SRp20 transactivate splicing through different and specific enhancers. RNA **5,** 468-483

Cazalla, D., Zhu, J., Manche, L., Huber, E., Krainer, A.R. and Cáceres, J.F. (2002) Nuclear export and retention signals in the RS domain of SR proteins. Mol Cell Biol. **22**, 6871-6882

Cazalla, D., Newton, K. and Cáceres, J.F. (2005) A novel SR-related protein is required for the second step of Pre-mRNA splicing. Mol Cell Biol. **25**, 2969-2980

Chabot, B., Blanchette, M., Lapierre, I. and La Branche, H. (1997) An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1. Mol Cell Biol. **17**, 1776-1786

Chan, C.B., Liu, X., Tang, X., Fu, H. and Ye, K. (2007) Akt phosphorylation of zyxin mediates its interaction with acinus-S and prevents acinus-triggered chromatin condensation. Cell Death Differ. **14**,1688-1699

Chang, Y.F., Imam, J.S. and Wilkinson, M.F. (2007) The nonsense-mediated decay RNA surveillance pathway. Annu Rev Biochem. **76**, 51-74

Chaudhary, N., McMahon, C. and Blobel, G. (1991) Primary structure of a human arginine-rich nuclear protein that colocalizes with spliceosome components. Proc Natl Acad Sci U S A. **88**, 8189-8193

Cheng, J., Belgrader, P., Zhou, X. and Maquat, L.E. (1994) Introns are cis effectors of the nonsense-codon-mediated reduction in nuclear mRNA abundance. Mol Cell Biol. **14**, 6317-6325

Cheung, W.L., Ajiro, K., Samejima, K., Kloc, M., Cheung, P., Mizzen, C.A., Beeser, A., Etkin, L.D., Chernoff, J., Earnshaw, W.C. and Allis, C.D. (2003) Apoptotic phosphorylation of histone H2B is mediated by mammalian sterile twenty kinase. Cell **113**, 507-517

Chou, T.B., Zachar, Z. and Bingham, P.M. (1987) Developmental expression of a regulatory gene is programmed at the level of splicing. EMBO J. **6,** 4095-4104

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

Chusainow, J., Ajuh, P.M., Trinkle-Mulcahy, L., Sleeman, J.E., Ellenberg, J. and Lamond, A.I. (2005) FRET analyses of the U2AF complex localize the U2AF35/U2AF65 interaction in vivo and reveal a novel self-interaction of U2AF35. RNA **11,** 1201-1214

Cobianchi, F., Calvio, C., Stoppini, M., Buvoli, M. and Riva, S. (1993) Phosphorylation of human hnRNP protein A1 abrogates in vitro strand annealing activity. Nucleic Acids Res. **21**, 949-955

Collier,B., Goobar-Larsson,L., Sokolowski,M., and Schwartz,S. (1998). Translational inhibition in vitro of human papillomavirus type 16 L2 mRNA mediated through interaction with heterogenous ribonucleoprotein K and poly(rC)-binding proteins 1 and 2. J. Biol. Chem., **273,** 22648-22656

Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J.L., Bell, J.C. and Duncan, P.I. (1996) The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. EMBO J. **15,** 265-275

Cooper, T. A. and Mattox, W. (1997) The regulation of splice-site selection, and its role in human disease. Am. J. Hum. Genet., **61**, 259-266.

Corden, J. L. and Patturajan, M. (1997) A CTD function linking transcription to splicing. Trends Biochem. Sci., **22**, 413-416.

Cortes, J.J., Sontheimer, E.J., Seiwert, S.D., Steitz, J.A. (1993) Mutations in the conserved loop of human U5 snRNA generate use of novel cryptic 5' splice sites in vivo. EMBO J., **12**, 5181-5189.

Cote,C.A., Gautreau,D., Denegre,J.M., Kress,T.L., Terry,N.A., and Mowry,K.L. (1999). A Xenopus protein related to hnRNP I has a role in cytoplasmic RNA localization. Mol. Cell **4,** 431-437

Coulter, L.R., Landree, M.A. and Cooper, T.A. (1997) Identification of a new class of exonic splicing enhancers by in vivo selection. Mol.Cell Biol. **17,** 2143-2150

Cowper, A.E., Caceres, J.F., Mayeda, A. and Screaton, G.R. (2001) Serine-Arginine (SR) Protein-like Factors That Antagonize Authentic SR Proteins and Regulate Alternative Splicing. J.Biol.Chem. **276,** 48908-48914

Cramer, P., Bushnell, D. A. and Kornberg R.D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science, **292**, 1863-1876.

Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. Genome Res. **14**, 1188-1190

Das, R., Yu, J., Zhang, Z., Gygi, M.P., Krainer, A.R., Gygi, S.P. and Reed, R. (2007) SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. Mol.Cell **26,** 867-881

Das, R., Zhou, Z. and Reed, R. (2000) Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. Mol Cell **5**, 779-787

Daugas, E., Susin, S.A., Zamzami, N., Ferri, K.F., Irinopoulou, T., Larochette, N., Prévost, M.C., Leber, B., Andrews, D., Penninger, J. and Kroemer, G. (2000) Mitochondrio-nuclear translocation of AIF in apoptosis and necrosis. FASEB J. **14**, 729-739

Dauksaite, V. and Akusjarvi, G. (2002) Human splicing factor ASF/SF2 encodes for a repressor domain required for its inhibitory activity on pre-mRNA splicing. J.Biol.Chem. **277,** 12579-12586

Dauksaite, V. and Akusjarvi, G. (2004) The second RNA-binding domain of the human splicing factor ASF/SF2 is the critical domain controlling adenovirus E1A alternative 5'-splice site selection. Biochem.J. **381,** 343-350

Dauwalder, B., Amaya-Manzanares, F. and Mattox, W. (1996) A human homologue of the Drosophila sex determination factor transformer-2 has conserved splicing regulatory functions. Proc.Natl.Acad.Sci.U.S.A **93,** 9004-9009

de la Grange, P., Dutertre, M., Martin, N. and Auboeuf, D. (2005) FAST DB: a website resource for the study of the expression regulation of human gene products. Nucleic Acids Res. **33**, 4276-4284

de la Mata, M. and Kornblihtt, A.R. (2006) RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. Nat.Struct.Mol.Biol. **13**, 973-980

de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D. and Kornblihtt, A.R. (2003) A Slow RNA Polymerase II Affects Alternative Splicing In Vivo. Mol.Cell **12,** 525-532

Deas, O., Dumont, C., MacFarlane, M., Rouleau, M., Hebib, C., Harper, F., Hirsch, F., Charpentier, B., Cohen, G.M., and Senik, A. (1998) Caspase-independent cell death induced by anti-CD2 or staurosporine in activated human peripheral T lymphocytes. J. Immunol. **161**, 3375-3383

Ding, J.H., Xu, X., Yang, D., Chu, P.H., Dalton, N.D., Ye, Z., Yeakley, J.M., Cheng, H., Xiao, R.P., Ross, J., Chen, J. and Fu, X.D. (2004) Dilated cardiomyopathy caused by tissue-specific ablation of SC35 in the heart. EMBO J. **23,** 885-896

Ding, J.H., Zhong, X.Y., Hagopian, J.C., Cruz, M.M., Ghosh, G., Feramisco, J., Adams, J.A. and Fu, X.D. (2006) Regulated cellular partitioning of SR protein-specific kinases in mammalian cells. Mol.Biol.Cell **17,** 876-885

Dostie, J. and Dreyfuss, G. (2002) Translation is required to remove Y14 from mRNAs in the cytoplasm. Curr Biol. **12**, 1060-1067

Dreyfuss, G., Matunis, M.J., Pinol-Roma, S. and Burd, C.G. (1993) hnRNP proteins and the biogenesis of mRNA. Annu.Rev.Biochem. **62,** 289-321

Earnshaw, W. C., Martins, L. M. and Kaufmann, S.H. (1999) Mammalian caspases: structure, activation, substrates, and functions during apoptosis. Annu. Rev. Biochem., **68**, 383-424.

Eldridge, A.G., Li, Y., Sharp, P.A. and Blencowe, B.J. (1999) The SRm160/300 splicing coactivator is required for exon-enhancer function. Proc.Natl.Acad.Sci.U.S.A **96,** 6125-6130

Ellis, J.D., Lleres, D., Denegri, M., Lamond, A.I. and Caceres, J.F. (2008) Spatial mapping of splicing factor complexes involved in exon and intron definition. J.Cell Biol. **181,** 921-934

Enari, M., Sakahira, H., Yokoyoma, H., Okawa, K., Iwamtsu, A., and Nagata, S. (1998) A caspase-activated DNase that degrades DNA during apoptosis, and its inhibitor ICAD. Nature **391**, 43-50

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

Eperon, I.C., Makarova, O.V., Mayeda, A., Munroe, S.H., Caceres, J.F., Hayward, D.G. and Krainer, A.R. (2000) Selection of alternative 5' splice sites: role of U1 snRNP and models for the antagonistic effects of SF2/ASF and hnRNP A1. Mol.Cell Biol. **20,** 8303-8318

Exline, C.M., Feng, Z. and Stoltzfus, C.M. (2008) Negative and positive mRNA splicing elements act competitively to regulate human immunodeficiency virus type 1 vif gene expression. J.Virol. **82,** 3921-3931

Fairbrother, W.G., Yeh, R.F., Sharp, P.A. and Burge, C.B. (2002) Predictive identification of exonic splicing enhancers in human genes. Science **297,** 1007-1013

Fic, W., Juge, F., Soret, J. and Tazi, J. (2007) Eye Development under the control of SRp55/B52-Mediated Alternative Splicing of eyeless. PLoS.ONE. **2,** e253

Fischer, D.C., Noack, K., Runnebaum, I.B., Watermann, D.O., Kieback, D.G., Stamm, S. and Stickeler, E. (2004) Expression of splicing factors in human ovarian cancer. Oncol.Rep. **11,** 1085-1090

Franke, T.F. (2008) Intracellular signaling by Akt: bound to be specific. Sci Signal. **1,** pe29

Fu, X.D. (1995) The superfamily of arginine/serine-rich splicing factors. RNA **1,** 663-680

Fu, X.D. and Maniatis, T. (1990) Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. Nature **343,** 437-441

Fu, X.D., Mayeda, A., Maniatis, T. and Krainer, A.R. (1992) General splicing factors SF2 and SC35 have equivalent activities in vitro, and both affect alternative 5' and 3' splice site selection. Proc.Natl.Acad.Sci.U.S.A **89,** 11224-11228

Gallego, M.E., Gattoni, R., Stevenin, J., Marie, J. and Expert-Bezancon, A. (1997) The SR splicing factors ASF/SF2 and SC35 have antagonistic effects on intronic enhancer-dependent splicing of the beta-tropomyosin alternative exon 6A. EMBO J. **16,** 1772-1784.

Garriga, J. and Grana, X. (2004) Cellular control of gene expression by T-type cyclin/CDK9 complexes. Gene **337,** 15-23

Ge, H. and Manley, J.L. (1990) A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. Cell **62,** 25-34

Ge, H., Zuo, P. and Manley, J.L. (1991) Primary structure of the human splicing factor ASF reveals similarities with Drosophila regulators. Cell **66,** 373-382

Giorgi, C. and Moore, M.J. (2007) The nuclear nurture and cytoplasmic nature of localized mRNPs. Semin Cell Dev Biol. **18**, 186-193

Gozani, O., Feld, R. and Reed, R. (1996) Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. Genes Dev., **10**, 233-243.

Gozani, O., Potashkin, J. and Reed, R. (1998) A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. Mol. Cell Biol., **18**, 4752-4760.

Graveley, B.R. (2000) Sorting out the complexity of SR protein functions. RNA **6,** 1197-1211

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

Graveley, B.R., Hertel, K.J. and Maniatis, T. (2001) The role of U2AF35 and U2AF65 in enhancer-dependent splicing. RNA **7,** 806-818

Green, D. R. and Kroemer, G. (2004) The pathophysiology of mitochondrial cell death. Science, **305**, 626-629.

Gui, J.F., Lane, W.S. and Fu, X.D. (1994) A serine kinase regulates intracellular localization of splicing factors in the cell cycle. Nature **369,** 678-682

Guil, S. and Caceres, J.F. (2007) The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. Nat Struct Mol Biol. **14**, 591-596

Guil, S., Long, J.C. and Cáceres J.F. (2006) hnRNP A1 relocalization to the stress granules reflects a role in the stress response. Mol Cell Biol. **26**, 5744-5758

Habelhah,H., Shah,K., Huang,L., Ostareck-Lederer,A., Burlingame,A.L., Shokat,K.M., Hentze,M.W., and Ronai,Z. (2001). ERK phosphorylation drives cytoplasmic accumulation of hnRNP-K and inhibition of mRNA translation. Nat. Cell Biol., **3,** 325-330.

Hacker, G. (2000) The morphology of apoptosis. Cell Tissue Res. **301**, 5-17

Hamilton, B.J., Burns, C.M., Nichols, R.C. and Rigby, W.F. (1997) Modulation of AUUUA response element binding by heterogeneous nuclear ribonucleoprotein A1 in human T lymphocytes. The roles of cytoplasmic location, transcription, and phosphorylation. J Biol Chem. **272**, 28732-28741

Hanamura, A., Caceres, J.F., Mayeda, A., Franza, B.R. and Krainer, A.R. (1998) Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors. RNA **4,** 430-444

Hastings, M.L. and Krainer, A.R. (2001) Functions of SR proteins in the U12-dependent AT-AC pre-mRNA splicing pathway. RNA **7,** 471-482

Hedley, M.L., Amrein, H. and Maniatis, T. (1995) An amino acid sequence motif sufficient for subnuclear localization of an arginine/serine-rich splicing factor. Proc Natl Acad Sci U S A.**92**, 11524-11528

Henne, W.M., Oomman, S., Attridge, J., Finckbone, V., Coates, P., Bliss, R., Strahlendorf, H. and Strahlendorf, J. (2006) AMPA-induced excitotoxicity increases nuclear levels of CAD, endonuclease G, and acinus and induces chromatin condensation in rat hippocampal pyramidal neurons. Cell Mol Neurobiol. **26**, 321-339

Hertel, K.J. and Graveley, B.R. (2005) RS domains contact the pre-mRNA throughout spliceosome assembly. Trends Biochem.Sci. **30,** 115-118

Hicks, M.J., Yang, C.R., Kotlajich, M.V. and Hertel, K.J. (2006) Linking splicing to Pol II transcription stabilizes pre-mRNAs and influences splicing patterns. PLoS.Biol. **4,** e147

Hillman, R.T., Green, R.E. and Brenner, S.E. (2004) An unappreciated role for RNA surveillance. Genome Biol. **5**, R8

Hirose, Y., Tacke, R. and Manley, J.L. (1999) Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. Genes Dev., **13**, 1234-1239.

Hirsch, T., Marchetti, P., Susin, S.A., Dallaporta, B., Zamzami, N., Marzo, I., Geuskens, M., and Kroemer, G. (1997) The apoptosis-necrosis paradox. Apoptogenic proteases activated after mitochondrial permeability transition determine the mode of cell death. Oncogene **15**, 1573-1582

Hoek,K.S., Kidd,G.J., Carson,J.H., and Smith,R. (1998). hnRNP A2 selectively binds the cytoplasmic transport sequence of myelin basic protein mRNA. Biochemistry, **37,** 7021-7029.

Hoffman, B.E. and Lis, J.T. (2000) Pre-mRNA splicing by the essential Drosophila protein B52: tissue and target specificity. Mol.Cell Biol. **20,** 181-186

Holz, M.K., Ballif, B.A., Gygi, S.P. and Blenis, J. (2005) mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events. Cell **123,** 569-580

Horiuchi, S., Koyanagi, Y., Zhou, Y., Miyamoto, H., Tanaka, Y., Waki, M., Matsumoto, A., Yamamoto, M. and Yamamoto, N. (1994) Soluble interleukin-6 receptors released from T cell or granulocyte/macrophage cell lines and human peripheral blood mononuclear cells are generated through an alternative splicing mechanism. Eur J Immunol. **24**, 1945-1948

Hromas, R., Boswell, S., Shen, R.N., Burgess, G., Davidson, A., Cornetta, K., Sutton, J. and Robertson, K. (1996) Forced over-expression of the myeloid zinc finger gene MZF-1 inhibits apoptosis and promotes oncogenesis in interleukin-3-dependent FDCP.1 cells. Leukemia **10**, 1049-1050

Hu, Y., Liu, Z., Yang, S.J. and Ye, K. (2007) Acinus-provoked protein kinase C delta isoform activation is essential for apoptotic chromatin condensation. Cell Death Differ. **14**, 2035-2046

Hu, Y., Yao, J., Liu, Z., Liu, X., Fu, H. and Ye, K. (2005) Akt phosphorylates acinus and inhibits its proteolytic cleavage, preventing chromatin condensation. EMBO J. **24**, 3543-3554

Hua, Y., Vickers, T.A., Okunola, H.L., Bennett, C.F. and Krainer, A.R. (2008) Antisense masking of an hnRNP A1/A2 intronic splicing silencer corrects SMN2 splicing in transgenic mice. Am.J.Hum.Genet. **82,** 834-848

Huang, Y. and Steitz, J.A. (2001) Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. Mol.Cell **7,** 899-905

Huang, Y. and Steitz, J.A. (2005) SRprises along a Messenger's Journey. Mol.Cell **17,** 613-615

Huang, Y., Gattoni, R., Stevenin, J. and Steitz, J.A. (2003) SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. Mol.Cell **11,** 837-843

Huang, Y., Yario, T.A. and Steitz, J.A. (2004) A molecular link between SR protein dephosphorylation and mRNA export. Proc.Natl.Acad.Sci.U.S.A **101,** 9666-9670

Hughes, T.A. (2006) Regulation of gene expression by alternative untranslated regions. Trends Genet. **22**, 119-122

Hutchinson, J.N., Ensminger, A.W., Clemson, C.M., Lynch, C.R., Lawrence, J.B. and Chess, A. (2007) A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. BMC Genomics **8,** 39

Ibrahim, E.C., Schaal, T.D., Hertel, K.J., Reed, R. and Maniatis, T. (2005) Serine/arginine-rich protein-dependent suppression of exon skipping by exonic splicing enhancers. Proc.Natl.Acad.Sci.U.S.A **102,** 5002-5007

Ihara, M., Stein, P. and Schultz, R.M. (2008) UBE2I (UBC9), a SUMO-Conjugating Enzyme, Localizes to Nuclear Speckles and Stimulates Transcription in Mouse Oocytes. Biol Reprod. Aug 13.

Inoue, M., Takahashi, K., Niide, O., Shibata, M., Fukuzawa, M. and Ra, C. (2005) LDOC1, a novel MZF-1-interacting protein, induces apoptosis. FEBS Lett. **579**, 604-608

Ishikawa, K., Nagase, T., Suyama, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O. (1998) The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro. DNA Res. **5**, 169-176

Izaurralde,E., Jarmolowski,A., Beisel,C., Mattaj,I.W., Dreyfuss,G., and Fischer,U. (1997). A role for the M9 transport signal of hnRNP A1 in mRNA nuclear export. J. Cell Biol., **137,** 27-35.

Jang, S.W., Yang, S.J., Ehlén, A., Dong, S., Khoury, H., Chen, J., Persson, J.L. and Ye, K. (2008) Serine/arginine protein-specific kinase 2 promotes leukemia cell proliferation by phosphorylating acinus and regulating cyclin A1. Cancer Res. **68**, 4559-4570

Jiang, Z. H., Zhang, W. J., Rao, Y. and Wu, J.Y. (1998) Regulation of Ich-1 pre-mRNA alternative splicing and apoptosis by mammalian splicing factors. Proc. Natl. Acad. Sci. U. S. A, **95**, 9155-9160.

Johnson, J.M., Castle, J., Garrett-Engele, P., Kan, Z., Loerch, P.M., Armour, C.D., Santos, R., Schadt, E.E., Stoughton, R. and Shoemaker, D.D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. Science **302,** 2141-2144

Joselin, A.P., Schulze-Osthoffe, K. and Schwerk, C. (2006) Loss of Acinus inhibits oligonuclesomal DNA fragmentation but not chromatin condensation during apoptosis. J. Biol. Chem.**281**, 12475-12484

Jumaa, H. and Nielsen, P.J. (1997) The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. EMBO J. **16,** 5077-5085

Jumaa, H., Wei, G. and Nielsen, P.J. (1999) Blastocyst formation is blocked in mouse embryos lacking the splicing factor SRp20. Curr.Biol. **9,** 899-902

Jurica, M.S. and Moore, M.J. (2003) Pre-mRNA Splicing. Awash in a Sea of Proteins. Mol.Cell **12,** 5-14

Kanopka, A., Muhlemann, O. and Akusjarvi, G. (1996) Inhibition by SR proteins of splicing of a regulated adenovirus pre- mRNA. Nature **381,** 535-538

Karni, R., de Stanchina, E., Lowe, S.W., Sinha, R., Mu, D. and Krainer, A.R. (2007) The gene encoding the splicing factor SF2/ASF is a proto-oncogene. Nat.Struct.Mol.Biol. **14,** 185-193

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G. and Ohno, S. (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**, 355-367

Kashima, T. and Manley, J.L. (2003) A negative element in SMN2 exon 7 inhibits splicing in spinal muscular atrophy. Nat.Genet. **34,** 460-463

Kashima, T., Rao, N., David, C.J. and Manley, J.L. (2007) hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. Hum Mol Genet. **16**, 3149-3159

Kataoka, N., Bachorik, J.L. and Dreyfuss, G. (1999) Transportin-SR, a nuclear import receptor for SR proteins. J.Cell Biol. **145,** 1145-1152

Kataoka, N., Yong, J., Kim, V.N., Velazquez, F., Perkinson, R.A., Wang, F. and Dreyfuss, G. (2000) Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. Mol Cell. **6**, 673-682

Kawano, T., Fujita, M. and Sakamoto, H. (2000) Unique and redundant functions of SR proteins, a conserved family of splicing factors, in caenorhabditis elegans development. Mech.Dev. **95,** 67-76

Kedersha, N., Cho, M.R., Li, W., Yacono, P.W., Chen, S., Gilks, N., Golan, D.E. and Anderson, P. (2000) Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. J Cell Biol. **151**, 1257-1268

Kedersha, N.L., Gupta, M., Li, W., Miller, I. and Anderson, P. (1999) RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. J Cell Biol. **147**, 1431-1442

Kedersha,N. and Anderson,P. (2002). Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. Biochem. Soc. Trans., **30,** 963-969.

Kennedy, C.F., Kramer, A. and Berget, S.M. (1998) A role for SRp54 during intron bridging of small introns with pyrimidine tracts upstream of the branch point. Mol.Cell Biol. **18,** 5425-5434

Kennedy, S.P., Weed, S.A., Forget, B.G. and Morrow, J.S. (1994) A partial structural repeat forms the heterodimer self-association site of all beta-spectrins. J Biol Chem. **269**, 11400-11408

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D. (2002) The human genome browser at UCSC. *Genome Res.* **12**, 996-1006

Kiledjian,M., DeMaria,C.T., Brewer,G., and Novick,K. (1997). Identification of AUF1 (heterogeneous nuclear ribonucleoprotein D) as a component of the alpha-globin mRNA stability complex. Mol. Cell Biol., **17,** 4870-4876.

Kim, S., Shi, H., Lee, D.K. and Lis, J.T. (2003) Specific SR protein-dependent splicing substrates identified through genomic SELEX. Nucleic Acids Res. **31,** 1955-1961

Kim, V.N., Kataoka, N. and Dreyfuss, G. (2001) Role of the nonsense-mediated decay factor hUpf3 in the splicing-dependent exon-exon junction complex. Science **293**, 1832-1836

Kohtz, J.D., Jamison, S.F., Will, C.L., Zuo, P., Luhrmann, R., Garcia-Blanco, M.A. and Manley, J.L. (1994) Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. Nature **368,** 119-124

Kornblihtt, A.R. (2007) Coupling transcription and alternative splicing. Adv.Exp.Med.Biol. **623,** 175-189

Krainer, A.R., Conway, G.C. and Kozak, D. (1990a) Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. Genes Dev. **4,** 1158-1171

Krainer, A.R., Conway, G.C. and Kozak, D. (1990b) The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. Cell **62,** 35-42

Krainer, A.R., Mayeda, A., Kozak, D. and Binns, G. (1991) Functional expression of cloned human splicing factor SF2: homology to RNA-binding proteins, U1 70K, and Drosophila splicing regulators. Cell **66,** 383-394

Kramer, A. (1996) The structure and function of proteins involved in mammalian pre-mRNA splicing. Annu.Rev.Biochem. **65,** 367-409

Krawczak, M., Reiss, J. and Cooper, D.N. (1992) The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Hum.Genet. **90,** 41-54

Kruhlak, M.J., Lever, M.A., Fischle, W., Verdin, E., Bazett-Jones, D.P. and Hendzel, M.J. (2000) Reduced mobility of the alternate splicing factor (ASF) through the nucleoplasm and steady state speckle compartments. J.Cell Biol. **150,** 41-51

Lai, M.C., Lin, R.I., Huang, S.Y., Tsai, C.W. and Tarn, W.Y. (2000) A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. J.Biol.Chem. **275,** 7950-7957

Lallena, M. J., Chalmers, K. J., Llamazares, S., Lamond, A.I. and Valcarcel, J. (2002) Splicing regulation at the second catalytic step by Sex-lethal involves 3' splice site recognition by SPF45. Cell, **109**, 285-296.

Lamond, A.I. and Spector, D.L. (2003) Nuclear speckles: a model for nuclear organelles. Nat.Rev.Mol.Cell Biol. **4,** 605-612

Lander, E. S., Linton, L. M. et al (2001) Initial sequencing and analysis of the human genome. Nature, **409**, 860-921

Lareau, L.F., Inada, M., Green, R.E., Wengrod, J.C. and Brenner, S.E. (2007) Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature **446**, 926-929

Le Hir, H. and Andersen, G.R. (2008) Structural insights into the exon junction complex. Curr Opin Struct Biol. **18**, 112-119

Le Hir, H., Gatfield, D., Izaurralde, E. and Moore, M.J. (2001) The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. EMBO J. **20**, 4987-4997

Le Hir, H., Izaurralde, E., Maquat, M.E. and Moore, M.J. (2000b) The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J. **19**, 6860-6869

Le Hir, H., Moore, M.J. and Maquat, M.E. (2000a) Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. Genes Dev. **14**, 1098-1108

Le Hir, H., Nott, A. and Moore, M.J. (2003) How introns influence and enhance eukaryotic gene expression. Trends Biochem Sci. **28**, 215-220

Lejeune, F., Ishigaki, Y., Li, X. and Maquat, L.E. (2002) The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. EMBO J. **21**, 3536-3545

Lewis, B.P., Green, R.E. and Brenner, S.E. (2003) Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc Natl Acad Sci U S A. **100**, 189-192

Li, H. and Bingham, P.M. (1991) Arginine/serine-rich domains of the su(wa) and tra RNA processing regulators target proteins to a subnuclear compartment implicated in splicing. Cell **67**, 335-342

Li, S. and Wilkinson, M.F. (1998) Nonsense surveillance in lymphocytes? Immunity **8**, 135-141

Li, X. and Manley, J.L. (2005) Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell **122,** 365-378

Li, X., Niu, T. and Manley, J.L. (2007) The RNA binding protein RNPS1 alleviates ASF/SF2 depletion-induced genomic instability. RNA **13,** 2108-2115

Lin, S. and Fu, X.D. (2007) SR proteins and related factors in alternative splicing. Adv.Exp.Med.Biol. **623,** 107-122

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S. and Fu, X.D. (2008) The splicing factor SC35 has an active role in transcriptional elongation. Nat.Struct.Mol.Biol., **15**, 819-826

Lin, S., Xiao, R., Sun, P., Xu, X. and Fu, X.D. (2005) Dephosphorylation-dependent sorting of SR splicing factors during mRNP maturation. Mol.Cell **20,** 413-425

Liu, H.X., Cartegni, L., Zhang, M.Q. and Krainer, A.R. (2001) A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. Nat.Genet. **27,** 55-58

Liu, H.X., Chew, S.L., Cartegni, L., Zhang, M.Q. and Krainer, A.R. (2000) Exonic splicing enhancer motif recognized by human SC35 under splicing conditions. Mol.Cell Biol. **20,** 1063-1071

Liu, H.X., Zhang, M. and Krainer, A.R. (1998) Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. Genes Dev. **12,** 1998-2012

Liu, X., Zou, H., Slaughter, C., and Wang, X. (1997) DFF, a heterodimeric protein that functions downstream of caspase 3 to trigger DNA fragmentation during apoptosis. Cell **89**, 175-184

Long, J.C. and Caceres, J.F. (2008) The SR protein family of splicing factors: master regulators of gene expression. Biochemical Journal. In press.

Longman, D., Johnstone, I.L. and Caceres, J.F. (2000) Functional characterization of SR and SR-related genes in Caenorhabditis elegans. EMBO J. **19,** 1625-1637

Longman, D., McGarvey, T., McCracken, S., Johnstone, I.L., Blencowe, B.J. and Caceres, J.F. (2001) Multiple interactions between SRm160 and SR family proteins in enhancer- dependent splicing and development of C. elegans. Curr.Biol. **11,** 1923-1933

Lorenz, C., von Pelchrzim, F. and Schroeder, R. (2006) Genomic systematic evolution of ligands by exponential enrichment (Genomic SELEX) for the identification of protein-binding RNAs independent of their expression levels. Nat.Protoc. **1,** 2204-2212

Los, M., Wesselborg, S. et al (1999) The role of caspases in development, immunity, and apoptotic signal transduction: lessons from knockout mice. Immunity., **10**, 629-639.

Lynch, K.W. (2007) Regulation of alternative splicing by signal transduction pathways. Adv.Exp.Med.Biol. **623,** 161-174

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

Makarova, O.V., Makarov, E.M. and Luhrmann, R. (2001) The 65 and 110 kDa SR-related proteins of the U4/U6.U5 tri-snRNP are essential for the assembly of mature spliceosomes. EMBO J. **20,** 2553-2563

Maniatis, T. and Reed, R. (2002) An extensive network of coupling among gene expression machines. Nature. **416**, 499-506

Maroney, P. A., Romfo, C. M. and Nilsen, T.W. (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**, 317-328.

Martinez-Contreras, R., Cloutier, P., Shkreta, L., Fisette, J.F., Revil, T. and Chabot, B. (2007) hnRNP proteins and splicing control. Adv.Exp.Med.Biol. **623,** 123-147

Matlin, A. J., Clark, F. and Smith, C.W. (2005) Understanding alternative splicing: towards a cellular code. Nat. Rev. Mol. Cell Biol., **6**, 386-398.

Matlin, A.J. and Moore, M.J. (2007) Spliceosome assembly and composition. Adv.Exp.Med.Biol. **623,** 14-35

Mayeda, A. and Krainer, A.R. (1992) Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell **68,** 365-375

Mayeda, A., Munroe, S.H., Cáceres, J.F. and Krainer, A.R. (1994) Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. EMBO J. **13**, 5483-5495

McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S. and Bentley, D.L. (1997a) 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. Genes Dev., **11**, 3306-3318.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M. and Bentley, D.L. (1997b) The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature, **385**, 357-361.

McGlincy, N.J. and Smith, C.W. (2008) Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? Trends Biochem Sci. **33**, 385-393

Melino, G., De, Laurenzi, V et al (2002) p73: Friend or foe in tumorigenesis. Nat. Rev. Cancer, **2**, 605-615.

Merdzhanova, G., Edmond, V., De Seranno, S., Van den Broeck, A., Corcos, L., Brambilla, C., Brambilla, E., Gazzeri, S. and Eymin, B. (2008) E2F1 controls alternative splicing pattern of genes involved in apoptosis through upregulation of the splicing factor SC35. Cell Death Differ. Sep 19.

Mermoud, J.E., Cohen, P.T and Lamond, A.I. (1994) Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. EMBO J. **13,** 5679-5688

Merz, C., Urlaub, H., Will, C.L., and Luhrmann, R. (2007) Protein composition of human mRNPs spliced in vitro and differential requirements for mRNP protein recruitment. RNA **13,** 116-128

Michael,W.M., Choi,M., and Dreyfuss,G. (1995). A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. Cell, **83,** 415-422.

Michlewski, G., Sanford, J.R. and Caceres, J.F. (2008) The Splicing Factor SF2/ASF Regulates Translation Initiation by Enhancing Phosphorylation of 4E-BP1. Mol.Cell **30,** 179-189

Mili, S. and Steitz, J.A. (2004) Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. RNA **10,** 1692-1694

Misteli, T., Caceres, J.F. and Spector, D.L. (1997) The dynamics of a pre-mRNA splicing factor in living cells. Nature **387,** 523-527

Modrek, B., Resch, A., Grasso, C. and Lee, C. (2001) Genome-wide detection of alternative splicing in expressed sequences of human genes. Nucleic Acids Res., **29**, 2850-2859.

Montzka, K.A. and Steitz, J.A. (1998) Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. Proc Natl Acad Sci U S A. **85**, 8885-8889

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

Moroy, T. and Heyd, F. (2007) The impact of alternative splicing in vivo: mouse models show the way. RNA **13,** 1155-1171

Morris, D. P. and Greenleaf, A. L. (2000) The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. J. Biol. Chem., **275**, 39935-39943.

Municio, M.M., Lozano, J., Sanchez, P., Moscat, J. and Diaz-Meco, M.T. (1995) Identification of heterogeneous ribonucleoprotein A1 as a novel substrate for protein kinase C zeta. J. Biol. Chem. **270**, 15885-15891

Muro, A.F., Caputi, M., Pariyarath, R., Pagani, F., Buratti, E. and Baralle, F.E. (1999) Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display. Mol.Cell Biol. **19,** 2657-2671

Nagy, E. and Maquat, L.E. (1998) A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. Trends Biochem Sci. **23**, 198-199

Newman, A. J. and Norman, C. (1992) U5 snRNA interacts with exon sequences at 5' and 3' splice sites. Cell, **68**, 743-754.

Newman, A. J., Teigelkamp, S. and Beggs, J.D. (1995) snRNA interactions at 5' and 3' splice sites monitored by photoactivated crosslinking in yeast spliceosomes. RNA., **1**, 968-980.

Ngo, J.C., Chakrabarti, S., Ding, J.H., Velazquez-Dones, A., Nolen, B., Aubol, B.E., Adams, J.A., Fu, X.D. and Ghosh, G. (2005) Interplay between SRPK and Clk/Sty kinases in phosphorylation of the splicing factor ASF/SF2 is regulated by a docking motif in ASF/SF2. Mol.Cell **20,** 77-89

Ngo, J.C., Giang, K., Chakrabarti, S., Ma, C.T., Huynh, N., Hagopian, J.C., Dorrestein, P.C., Fu, X.D., Adams, J.A. and Ghosh, G. (2008) A sliding docking interaction is essential for sequential and processive phosphorylation of an SR protein by SRPK1. Mol.Cell **29,** 563-576

Nguyen, V.T., Kiss, T., Michels, A.A. and Bensaude, O. (2001) 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. Nature **414**, 322-325

Nicholas, K.B., Nicholas H.B. Jr., and Deerfield, D.W. II. (1997) GeneDoc: Analysis and Visualization of Genetic Variation. EMBNEW.NEWS **4**:14

Nikolakaki, E., Drousou, V., Sanidas, I., Peidis, P., Papamarcaki, T., Iakoucheva, L.M. and Giannakouros, T. (2008) RNA association or phosphorylation of the RS domain prevents aggregation of RS domain-containing proteins. Biochim. Biophys. Acta. **1780**, 214-225

Niranjanakumari, S., Lasda, E., Brazas, R. and Garcia-Blanco, M.A. (2002) Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions in vivo. Methods **26,** 182-190

Nott, A., Le Hir, H. and Moore, M.J. (2004) Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Genes Dev. **18**, 210-122

Ohno, M. and Shimura, Y. (1996) A human RNA helicase-like protein, HRH1, facilitates nuclear export of spliced mRNA by releasing the RNA from the spliceosome. Genes Dev. **10,** 997-1007

O'Keefe, R.T., Mayeda, A., Sadowski, C.L., Krainer, A.R. and Spector, D.L. (1994) Disruption of pre-mRNA splicing in vivo results in reorganization of splicing factors. J Cell Biol. **124**, 249-260

Olson, S., Blanchette, M., Park, J., Savva, Y., Yeo, G.W., Yeakley, J.M., Rio, D.C. and Graveley, B.R. (2007) A regulator of Dscam mutually exclusive splicing fidelity. Nat.Struct.Mol.Biol. **14**, 1134-1140

Ostareck,D.H., Ostareck-Lederer,A., Wilm,M., Thiele,B.J., Mann,M., and Hentze,M.W. (1997). mRNA silencing in erythroid differentiation: hnRNP K and hnRNP E1 regulate 15-lipoxygenase translation from the 3' end. Cell, **89,** 597-606

Pan, Q., Shai, O., Misquitta, C., Zhang, W., Saltzman, A.L., Mohammad, N., Babak, T., Siu, H., Hughes, T.R., Morris, Q.D., Frey, B.J. and Blencowe, B.J. (2004) Revealing global regulatory features of Mammalian alternative splicing using a quantitative microarray platform. Mol.Cell **16,** 929-941

Patel, N.A., Kaneko, S., Apostolatos, H.S., Bae, S.S., Watson, J.E., Davidowitz, K., Chappell, D.S., Birnbaum, M.J., Cheng, J.Q. and Cooper, D.R. (2005) Molecular and Genetic Studies Imply Akt-mediated Signaling Promotes Protein Kinase C{beta}II Alternative Splicing via Phosphorylation of Serine/Arginine-rich Splicing Factor SRp40. J.Biol.Chem. **280,** 14302-14309

Pauler, F.M., Koerner, M.V. and Barlow D.P. (2007) Silencing by imprinted noncoding RNAs: is transcription the answer? Trends Genet. **23**, 284-292

Peng, X. and Mount, S.M. (1995) Genetic enhancement of RNA-processing defects by a dominant mutation in B52, the Drosophila gene for an SR protein splicing factor. Mol.Cell Biol. **15,** 6273-6282

Pesole, G., Mignone, F., Gissi, C., Grillo, G., Licciulli, F. and Liuni, S. (2001) Structural and functional features of eukaryotic mRNA untranslated regions. Gene **276**, 73-81

Peterlin, B.M. and Price, D.H. (2006) Controlling the elongation phase of transcription with P-TEFb. Mol Cell **23**, 297-305

Phair, R.D. and Misteli, T. (2000) High mobility of proteins in the mammalian cell nucleus. Nature **404,** 604-609

Pino, I., Pío, R., Toledo, G., Zabalegui, N., Vicent, S., Rey, N., Lozano, M.D., Torre, W., García-Foncillas, J. and Montuenga, L.M. (2003) Altered patterns of expression of members of the heterogeneous nuclear ribonucleoprotein (hnRNP) family in lung cancer. Lung Cancer **41**, 131-143

Pinol-Roma,S. and Dreyfuss,G. (1992). Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. Nature **355,** 730-732.

Pollard,V.W., Michael,W.M., Nakielny,S., Siomi,M.C., Wang,F., and Dreyfuss,G. (1996). A novel receptor-mediated nuclear protein import pathway. Cell, **86,** 985-994.

Proc Natl Acad Sci U S A. **100**, 189-192

Quignon, F., DeBels, F., Koken, M., Feunteun, J., Ameisen, J.C., and de Thé, H. (1998) PML induces a novel caspase-independent death process. Nat. Gen. **20**, 259-265

Ramchatesingh, J., Zahler, A.M., Neugebauer, K.M., Roth, M.B. and Cooper, T.A. (1995) A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. Mol.Cell Biol. **15,** 4898-4907

Rappsilber, J., Ryder, U., Lamond, A.I. and Mann, M. (2002) Large-scale proteomic analysis of the human spliceosome. Genome Res. **12,** 1231-1245

Rasheva, V.I., Knight, D., Bozko, P., Marsh, K. and Frolov, M.V. (2006) Specific role of the SR protein splicing factor B52 in cell cycle control in Drosophila. Mol.Cell Biol. **26,** 3468-3477

Rathert, P., Dhayalan, A., Murakami, M., Zhang, X., Tamas, R., Jurkowska, R., Komatsu, Y., Shinkai, Y., Cheng, X. and Jeltsch, A. (2008) Protein lysine methyltransferase G9a acts on non-histone targets. Nat Chem Biol. **4**, 344-346

Rebane,A., Aab,A., and Steitz,J.A. (2004). Transportins 1 and 2 are redundant nuclear import factors for hnRNP A1 and HuR. RNA., **10,** 590-599.

Reichert, V.L., Le Hir, H., Jurica, M.S. and Moore, M.J. (2002) 5' exon interactions within the human spliceosome establish a framework for exon junction complex structure and assembly. Genes Dev. **16**, 2778-2791

Ring, H.Z. and Lis, J.T. (1994) The SR protein B52/SRp55 is essential for Drosophila development. Mol.Cell Biol. **14,** 7499-7506

Rino, J., Desterro, J.M., Pacheco, T.R., Gadella, T.W., Jr. and Carmo-Fonseca, M. (2008) Splicing factors SF1 and U2AF associate in extra-spliceosomal complexes. Mol.Cell Biol. **28**, 3045-3057

Robberson, B.L., Cote, G.J. and Berget, S.M. (1990) Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol.Cell Biol. **10,** 84-94

Roberts, G.C. and Smith, C.W. (2002) Alternative splicing: combinatorial output from the genome. Curr Opin Chem Biol. **6**, 375-383

Robertson, K.A., Hill, D.P., Kelley, M.R., Tritt, R., Crum, B., Van Epps, S., Srour, E., Rice, S. and Hromas, R. (1998) The myeloid zinc finger gene (MZF-1) delays retinoic acid-induced apoptosis and differentiation in myeloid leukemia cells. Leukemia **12**, 690-698

Rooke, N., Markovtsov, V., Cagavi, E. and Black, D.L. (2003) Roles for SR Proteins and hnRNP A1 in the Regulation of c-src Exon N1. Mol.Cell Biol. **23,** 1874-1884

Roscigno, R.F. and Garcia-Blanco, M.A. (1995) SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. RNA **1,** 692-706

Rossi, F., Labourier, E., Forne, T., Divita, G., Derancourt, J., Riou, J.F., Antoine, E., Cathala, G., Brunel, C. and Tazi, J. (1996) Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. Nature **381,** 80-82

Roth, M.B., Murphy, C. and Gall, J.G. (1990) A monoclonal antibody that recognizes a phosphorylated epitope stains lampbrush chromosome loops and small granules in the amphibian germinal vesicle. J.Cell Biol. **111,** 2217-2223

Sahara, S., Aoto, M. et al (1999) Acinus is a caspase-3-activated protein required for apoptotic chromatin condensation. Nature, **401**, 168-173.

Sakahira, H., Enari, M., and Nagata, S. (1998) Cleavage of CAD inhibitor in CAD activation and DNA degradation during apoptosis. Nature **391**, 96-99

Sambrook, J., Fritsch, E.F. and Maniatis, T.(1989) Molecular Cloning: A Laboratory Manual. Second Edition. Cold Spring Harbour Laboratory Press.

Samejima, K., Toné, S., Kottke, T.J., Enari, M., Sakahira, H., Cooke, C.A., Durrieu, F., Martins, L.M., Nagata, S., Kaufmann, S.H., and Earnshaw, W.C. (1998) Transition from caspase-dependent to caspase-independent mechanisms at the onset of apoptotic execution. J. Cell Biol. **143**, 225-239

Sanford, J.R. and Bruzik, J.P. (1999) Developmental regulation of SR protein phosphorylation and activity. Genes Dev. **13,** 1513-1518

Sanford, J.R., Coutinho, P., Hackett, J.A., Wang, X., Ranahan, W. and Caceres,J.F. (2008) Identification of nuclear and cytoplasmic mRNA targets for the shuttling protein SF2/ASF. PLoS ONE **3**, e3369

Sanford, J.R., Ellis, J.D., Cazalla, D. and Caceres, J.F. (2005) Reversible phosphorylation differentially affects nuclear and cytoplasmic functions of splicing factor 2/alternative splicing factor. Proc.Natl.Acad.Sci.U.S.A **102,** 15042-15047

Sanford, J.R., Gray, N.K., Beckmann, K. and Caceres, J.F. (2004) A novel role for shuttling SR proteins in mRNA translation. Genes Dev. **18,** 755-768

Sato, H., Hosoda, N. and Maquat, L.E. (2008) Efficiency of the pioneer round of translation affects the cellular site of nonsense-mediated mRNA decay. Mol.Cell **29,** 255-262

Schaal, T.D. and Maniatis, T. (1999) Selection and characterization of pre-mRNA splicing enhancers: identification of novel SR protein-specific enhancer sequences. Mol.Cell Biol. **19,** 1705-1719

Schmucker, D., Clemens, J.C., Shu, H., Worby, C.A., Xiao, J., Muda, M., Dixon, J.E. and Zipursky, S.L. (2000)..Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell **101**, 671-84

Schwerk, C. and Schulze-Osthoff, K. (2003) Non-apoptotic functions of caspases in cellular proliferation and differentiation. Biochem. Pharmacol., **66**, 1453-1458.

Schwerk, C. and Schulze-Osthoff, K. (2005) Regulation of apoptosis by alternative pre-mRNA splicing. Mol. Cell, **19**, 1-13.

Schwerk, C., Prasad, J., Degenhardt, K., Erdjument-Bromage, H., White, E., Tempst, P., Kidd, V.J., Manley, J.L., Lahti, J.M. and Reinberg, D. (2003) ASAP, a novel protein complex involved in RNA processing and apoptosis. Mol. Cell Biol., **23**, 2981-2990.

Screaton, G.R., Caceres, J.F., Mayeda, A., Bell, M.V., Plebanski, M., Jackson, D.G., Bell, J.I. and Krainer, A.R. (1995) Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. EMBO J. **14,** 4336-4349

Shaham, S. and Horvitz, H. R. (1996) An alternatively spliced C. elegans ced-4 RNA encodes a novel cell death inhibitor. Cell, **86**, 201-208.

Sharp, P. A. (1994) Split genes and RNA splicing. Cell, **77**, 805-815.

Sharp, P. A. and Burge, C. B. (1997) Classification of introns: U2-type or U12-type. Cell, **91**, 875-879.

Shen, H. and Green, M.R. (2004) A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly. Mol.Cell **16,** 363-373

Shen, H. and Green, M.R. (2007) RS domain-splicing signal interactions in splicing of U12-type and U2-type introns. Nat.Struct.Mol.Biol. **14,** 597-603

Shen, H., Kan, J.L. and Green, M.R. (2004) Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. Mol.Cell **13,** 367-376

Shepard, P.J. and Hertel, K.J. (2008) Conserved RNA secondary structures promote alternative splicing. RNA **14**, 1463-1469

Shi, H., Hoffman, B.E. and Lis, J.T. (1997) A specific RNA hairpin loop structure binds the RNA recognition motifs of the Drosophila SR protein B52. Mol.Cell Biol. **17,** 2649-2657

Shi, J., Hu, Z., Pabon, K. and Scotto, K.W. (2008) Caffeine regulates alternative splicing in a subset of cancer-associated genes: a role for SC35. Mol.Cell Biol. **28,** 883-895

Shi, Y. (2002) Apoptosome: the cellular engine for the activation of caspase-9. Structure., **10**, 285-288.

Shi, Y. and Manley, J.L. (2007) A complex signaling pathway regulates SRp38 phosphorylation and pre-mRNA splicing in response to heat shock. Mol.Cell **28,** 79-90

Shibuya, T., Tange, T.O., Sonenberg, N. and Moore, M.J. (2004) eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. Nat Struct Mol Biol. **11**, 346-351

Shin, C., Feng, Y. and Manley, J.L. (2004) Dephosphorylated SRp38 acts as a splicing repressor in response to heat shock. Nature **427,** 553-558

Shomron, N., Alberstein, M., Reznik, M. and Ast, G. (2005) Stress alters the subcellular distribution of hSlu7 and thus modulates alternative splicing. J Cell Sci. **118**, 1151-1159

Simard, M.J. and Chabot, B. (2002) SRp30c is a repressor of 3' splice site utilization. Mol.Cell Biol. **22,** 4001-4010

Singer, B.S., Shtatland, T., Brown, D. and Gold, L. (1997) Libraries for genomic SELEX. Nucleic Acids Res. **25,** 781-786

Siomi,H. and Dreyfuss,G. (1995). A nuclear localization domain in the hnRNP A1 protein. J. Cell Biol., **129,** 551-560.

Smith, C.W. and Valcarcel, J. (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem.Sci. **25,** 381-388

Smith, E.R., Cayrou, C., Huang, R., Lane, W.S., Côté, J. and Lucchesi, J.C. (2005) A human protein complex homologous to the Drosophila MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. Mol Cell Biol. **25**, 9175-9188

Solis, A.S., Peng, R., Crawford, J.B., Phillips, J.A., III and Patton, J.G. (2008) Growth Hormone Deficiency and Splicing Fidelity: Two Serine/Arginine-rich proteins, ASF/SF2 and SC35, act antagonistically. J.Biol.Chem., **283,** 23619-23626.

Sontheimer, E. J. and Steitz, J. A. (1992) Three novel functional variants of human U5 small nuclear RNA. Mol. Cell Biol., **12**, 734-746.

Spector, D.H., Schrier, W.H. and Busch, H. (1983) Immunoelectron microscopic localization of snRNPs. Biol Cell **49**, 1-10

Spritz, R.A., Strunk, K., Surowy, C.S., Hoch, S.O., Barton, D.E. and Francke, U. (1987) The human U1-70K snRNP protein: cDNA cloning, chromosomal localization, expression, alternative splicing and RNA-binding. Nucleic Acids Res. **15,** 10373-10391

Stickeler, E., Kittrell, F., Medina, D. and Berget, S.M. (1999) Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis. Oncogene **18,** 3574-3582

Sun, J. S. and Manley, J. L. (1995) A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. Genes Dev., **9**, 843-854.

Sun, X., Alzhanova-Ericsson, A.T., Visa, N., Aissouni, Y., Zhao, J. and Daneholt, B. (1998) The hrp23 protein in the balbiani ring pre-mRNP particles is released just before or at the binding of the particles to the nuclear pore complex. J Cell Biol. **142**, 1181-1193

Susin, S.A., Lorenzo, H.K., Zamzami, N., Marzo, I., Snow, B.E., Brothers, G.M., Mangion, J., Jacotot, E., Costantini, P., and Loeffler, M. (1999) Molecular characterization of mitochondrial apoptosis-inducing factor. Nature **397**, 441-446

Swartz, J.E., Bor, Y.C., Misawa, Y., Rekosh, D. and Hammarskjold, M.L. (2007) The shuttling SR protein 9G8 plays a role in translation of unspliced mRNA containing a constitutive transport element. J.Biol.Chem. **282,** 19844-19853

Sykes, S.M., Mellert, H.S., Holbert, M.A., Li, K., Marmorstein, R., Lane, W.S. and McMahon, S.B. (2006) Acetylation of the p53 DNA-binding domain regulates apoptosis induction. Mol Cell **24**, 841-851

Szymczyna, B.R., Bowman, J., McCracken, S., Pineda-Lucena, A., Lu, Y., Cox, B., Lambermon, M., Graveley, B.R., Arrowsmith, C.H. and Blencowe, B.J. (2003) Structure and function of the PWI motif: a novel nucleic acid-binding domain that facilitates pre-mRNA processing. Genes Dev. **17,** 461-475

Tacke, R. and Manley, J.L. (1995) The human splicing factors ASF/SF2 and SC35 possess distinct, functionally significant RNA binding specificities. EMBO J. **14,** 3540-3551

Tacke, R., Chen, Y. and Manley, J.L. (1997) Sequence-specific RNA binding by an SR protein requires RS domain phosphorylation: creation of an SRp40-specific splicing enhancer. Proc.Natl.Acad.Sci.U.S.A **94,** 1148-1153

Tacke, R., Tohyama, M., Ogawa, S. and Manley, J.L. (1998) Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing. Cell **93,** 139-148

Tange, T. O., Shibuya, T. et al (2005) Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. RNA., **11**, 1869-1883.

Tange, T.O., Nott, A. and Moore, M.J. (2004) The ever-increasing complexities of the exon junction complex. Curr.Opin.Cell Biol. **16,** 279-284

Tazi, J., Kornstadt, U., Rossi, F., Jeanteur, P., Cathala, G., Brunel, C. and Luhrmann, R. (1993) Thiophosphorylation of U1-70K protein inhibits pre-mRNA splicing. Nature **363,** 283-286

The ENCODE Project Consortium. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature **447**, 799–816

Theissen, H., Etzerodt, M., Reuter, R., Schneider, C., Lottspeich, F., Argos, P., Luhrmann, R. and Philipson, L. (1986) Cloning of the human cDNA for the U1 RNA-associated 70K protein. EMBO J. **5,** 3209-3217

Thermann, R., Neu-Yilik, G., Deters, A., Frede, U., Wehr, K., Hagemeier, C., Hentze, M.W. and Kulozik, A.E. (1998) Binary specification of nonsense codons by splicing and cytoplasmic translation. EMBO J. **17**, 3484-3494

Thomas, T., Dixon, M.P., Kueh, A.J. and Voss, A.K. (2008) Mof (MYST1 or KAT8) is essential for progression of embryonic development past the blastocyst stage and required for normal chromatin architecture. Mol Cell Biol. **28**, 5093-5105

Thomas, T., Loveland, K.L. and Voss, A.K. (2007) The genes coding for the MYST family histone acetyltransferases, Tip60 and Mof, are expressed at high levels during sperm development. Gene Expr Patterns **7**, 657-665

Thompson, C.B. (1995) Apoptosis in the pathogenesis and treatment of disease. Science **267**, 1456-1462

Tian, H. and Kole, R. (2001) Strong RNA splicing enhancers identified by a modified method of cycled selection interact with SR protein. J.Biol.Chem. **276,** 33833-33839

Tuerk, C. and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science **249,** 505-510

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A. and Darnell, R.B. (2003) CLIP identifies Nova-regulated RNA networks in the brain. Science **302,** 1212-1215

Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., Gaasterland, T., Blencowe, B.J. and Darnell, R.B. (2006) An RNA map predicting Nova-dependent splicing regulation. Nature **444,** 580-586

Ule, J., Ule, A., Spencer, J., Williams, A., Hu, J.S., Cline, M., Wang, H., Clark, T., Fraser, C., Ruggiu, M., Zeeberg, B.R., Kane, D., Weinstein, J.N., Blume, J. and Darnell, R.B. (2005) Nova regulates brain-specific splicing to shape the synapse. Nat.Genet. **37,** 844-852

Valadkhan, S. and Manley, J. L. (2001) Splicing-related catalysis by protein-free snRNAs. Nature, **413**, 701-707.

Valadkhan, S. and Manley, J. L. (2002) Intrinsic metal binding by a spliceosomal RNA. Nat. Struct. Biol., **9**, 498-499.

Valentine, C.R. (1998) The association of nonsense codons with exon skipping. Mutat.Res. **411,** 87-117

van der Houven van Oordt, Diaz-Meco,M.T., Lozano,J., Krainer,A.R., Moscat,J., and Caceres,J.F. (2000). The MKK(3/6)-p38-signaling cascade alters the subcellular distribution of hnRNP A1 and modulates alternative splicing regulation. J. Cell Biol., **149,** 307-316.

Van Herreweghe, E., Egloff, S., Goiffon, I., Jády, B.E., Froment, C., Monsarrat, B. and Kiss, T. (2007) Dynamic remodelling of human 7SK snRNP controls the nuclear level of active P-TEFb. EMBO J. **26**, 3570-3580

Venter, J. C., Adams, M. D. et al (2001) The sequence of the human genome. Science, **291**, 1304-1351.

Vousden, K. H. and Lu, X. (2002) Live or let die: the cell's response to p53. Nat. Rev. Cancer, **2**, 594-604.

Wagner, E. and Lykke-Andersen, J. (2002) mRNA surveillance: the perfect persist. J Cell Sci. **115**, 3033-3038

Wang, H.Y., Lin, W., Dyck, J.A., Yeakley, J.M., Songyang, Z., Cantley, L.C. and Fu, X.D. (1998) SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. J.Cell Biol. **140,** 737-750

Wang, H.Y., Xu, X., Ding, J.H., Bermingham, J.R., Jr. and Fu, X.D. (2001) SC35 plays a role in T cell development and alternative splicing of CD45. Mol.Cell **7,** 331-342

Wang, J. and Manley, J.L. (1995) Overexpression of the SR proteins ASF/SF2 and SC35 influences alternative splicing in vivo in diverse ways. RNA **1,** 335-346

Wang, J., Takagaki, Y. and Manley ,J.L. (1996) Targeted disruption of an essential vertebrate gene: ASF/SF2 is required for cell viability. Genes Dev. **10,** 2588-2599

Wang, L., Miura, M., Bergeron, L., Zhu, H. and Yuan, J. (1994) Ich-1, an Ice/ced-3-related gene, encodes both positive and negative regulators of programmed cell death. Cell, **78**, 739-750.

Wiegand, H.L., Lu, S. and Cullen, B.R. (2003) Exon junction complexes mediate the enhancing effect of splicing on mRNA expression. Proc. Natl. Acad. Sci. USA. **100**, 11327-11332

Will, C.L. and Luhrmann, R. (2001) Spliceosomal UsnRNP biogenesis, structure and function. Curr.Opin.Cell Biol. **13,** 290-301

Will, C.L. and Luhrmann, R. (2005) Splicing of a rare class of introns by the U12-dependent spliceosome. Biol.Chem. **386,** 713-724

Wirth, B., Brichta, L. and Hahnen, E. (2006) Spinal muscular atrophy: from gene to therapy. Semin.Pediatr.Neurol. **13,** 121-131

Wu, J.Y. and Maniatis, T. (1993) Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell **75,** 1061-1070

Wyatt, J. R., Sontheimer, E. J. and Steitz, J.A. (1992) Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. Genes Dev., **6**, 2542-2553.

Xiao, R., Sun, Y., Ding, J.H., Lin, S., Rose, D.W., Rosenfeld, M.G., Fu, X.D. and Li, X. (2007) Splicing regulator SC35 is essential for genomic stability and cell proliferation during mammalian organogenesis. Mol.Cell Biol. **27,** 5393-5402

Xiao, S.H. and Manley, J.L. (1997) Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. Genes Dev. **11,** 334-344

Xu, X., Yang, D., Ding, J.H., Wang, W., Chu, P.H., Dalton, N.D., Wang, H.Y., Bermingham, J.R., Jr., Ye, Z., Liu, F., Rosenfeld, M.G., Manley, J.L., Ross, J., Jr., Chen, J., Xiao, R.P., Cheng, H. and Fu, X.D. (2005) ASF/SF2-Regulated CaMKIIdelta Alternative Splicing Temporally Reprograms Excitation-Contraction Coupling in Cardiac Muscle. Cell **120,** 59-72

Xu,N., Chen,C.Y., and Shyu,A.B. (2001). Versatile role for hnRNP D isoforms in the differential regulation of cytoplasmic mRNA turnover. Mol. Cell Biol., **21,** 6960-6971.

Yang, X., Bani, M.R., Lu, S.J., Rowan, S., Ben David, Y. and Chabot, B. (1994) The A1 and A1B proteins of heterogeneous nuclear ribonucleoparticles modulate 5' splice site selection in vivo. Proc.Natl.Acad.Sci.U.S.A **91,** 6924-6928

Yang, Z., Zhu, Q., Luo, K. and Zhou, Q. (2001) The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. Nature **414**, 317-322

Yean, S. L., Wuenschell, G., Termini, J. and Lin, R.J. (2000) Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. Nature **408**, 881-884.

Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R.V., Gentile, C., Gebara, M. and Corden, J.L. (1996) The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. Proc.Natl.Acad.Sci.U.S.A **93,** 6975-6980

Zahler, A.M., Lane, W.S., Stolk, J.A. and Roth, M.B. (1992) SR proteins: a conserved family of pre-mRNA splicing factors. Genes Dev. **6,** 837-847

Zahler, A.M., Neugebauer, K.M., Lane, W.S. and Roth, M.B. (1993b) Distinct functions of SR proteins in alternative pre-mRNA splicing. Science **260,** 219-222

Zahler, A.M., Neugebauer, K.M., Stolk, J.A. and Roth, M.B. (1993a) Human SR proteins and isolation of a cDNA encoding SRp75. Mol Cell Biol. **13**, 4023-4028

Zech, V.F., Dlaska, M., Tzankov, A. and Hilbe, W. (2006) Prognostic and diagnostic relevance of hnRNP A2/B1, hnRNP B1 and S100 A2 in non-small cell lung cancer. Cancer Detect Prev. **30**, 395-402

Zhang, W.J. and Wu, J.Y. (1996) Functional properties of p54, a novel SR protein active in constitutive and alternative splicing. Mol.Cell Biol. **16,** 5400-5408

Zhang, X.H. and Chasin, L.A. (2004) Computational definition of sequence motifs governing constitutive exon splicing. Genes Dev. **18,** 1241-1250

Zhang, X.H., Kangsamaksin, T., Chao, M.S., Banerjee, J.K. and Chasin, L.A. (2005) Exon inclusion is dependent on predictable exonic splicing enhancers. Mol.Cell Biol. **25,** 7323-7332

Zhang, Z. and Krainer, A.R. (2004) Involvement of SR Proteins in mRNA Surveillance. Mol.Cell, **16** 597-607

Zhang, Z., Schwartz, S., Wagner, L. and Miller W. (2000) A greedy algorithm for aligning DNA sequences. J Comput Biol. **7**, 203-14

Zhou, Z. and Reed, R. (1998) Human homologs of yeast prp16 and prp17 reveal conservation of the mechanism for catalytic step II of pre-mRNA splicing. EMBO J. **17,** 2095-2106

Zhou, Z., Licklider, L.J., Gygi, S.P. and Reed, R. (2002) Comprehensive proteomic analysis of the human spliceosome. Nature **419,** 182-185

Zhu, J., Mayeda, A. and Krainer, A.R. (2001) Exon Identity Established through Differential Antagonism between Exonic Splicing Silencer-Bound hnRNP A1 and Enhancer-Bound SR Proteins. Mol.Cell **8,** 1351-1361

Appendices

CLIP id	Genomic coordinates	Ensembl gene id	id gene	Tag location	UTR type	Splicing Event
114	chr3:10,321,729-10,321,801	ENSG00000157020	SEC13	Exon 9		
482	chr3:10,321,729-10,321,801	ENSG00000157020	SEC13	Exon 9		
150	chr2:112,342,070-112,342,101	ENSG00000153107	ANAPC1	Exons 7 and 8		
72	chr3:15,098,953-15,098,991	ENSG00000131381	ZFYVE20	Exon 9		
186	chr11:205,813-205,834	ENSG00000142082	SIRT3	Exon 8	3'	
189	chr16:31,050,059-31,050,100	ENSG00000103510	MYST1	Exon 11	ORF and 3'	Intron 10 retention
198	chr13:72,229,403-72,230,096	ENSG00000083520	DIS3	Exon 23	3'	
209	chr12:97,433,766-97,433,786	ENSG00000120802	TMPO	Exon 1	5'	
212	chr16:1,299,711-1,299,736	ENSG00000103275	UBE2I	Exon 1	5'	Exon 2 skipping
227	chr3:52,698,204-52,698,246	ENSG00000163938	GNL3	Exon 6		Intron 7 retention
230	chr7:128,399,831-128,399,851	ENSG00000064419	TNPO3	Exon 20		Exon 23 skipping
261	chr19:4,743,870-4,743,893	ENSG00000141965	FEM1A	Exon 2		
272	chr2:182,774,700-182,774,718	ENSG00000115252	PDE1A	Exon11		
277	chr3:187,127,081-187,127,109	ENSG00000136527	SFRS10	Exon 4		Alt 5'ss exon 4
298	chr6:43,039,936-43,040,002	ENST00000304611	PEX6	Exon 17	3'	Internal exon 17 deletion
317	chr1:64,867,710-64,867,752	ENSG00000158966	CACHD1	Exon 8		Exon 7 and exon 9 skipping
364	chr16: 87422629-87422703	ENSG00000141012	GALNS	Exon 3		Intron 2 retention
383	chr11:62,149,412-62,149,442	ENSG00000089597	GANAB	Exon 25	3'	Internal exon 25 deletion
386	chr19:45,807,395-45,807,508	ENSG00000090006	LTBP4	Exons 15 and 16		
389	chr7:158,399,683-158,399,712	ENSG00000126870	WDR60	Exons 13 and 14		
395	chr5:80,752,036-80,752,070	ENSG00000145687	SSBP2	Exon 16	3'	
405	chr9:4,789,762-4,789,835	ENSG00000120158	RCL1	Intron 1		Exons 2 and 3 skipping
432	chr4:174,449,868-174,449,894	ENSG00000109586	GALNT7	Exon 3		
698	chr4:174,449,868-174,449,894	ENSG00000109586	GALNT7	Exon 3		
470	chr7:21,907,645-21,907,695	ENSG00000105877	DNAH11	Exon 82	3'	
486	chr1:144,746,396-144,746,504	ENSG00000152042	NBPF14	Exon 12		Exons 8-12 skipping
505	chr5:76,064,615-76,064,635	ENSG00000181104	F2R	Exon 2		

Appendix 1. A list of annotation of the CLIP tags identified for Acinus. Ensembl, UCSC Genome Browser and Fast-DB databases were used. Headers for the table are as follows: CLIP id: Id given to each tag during analysis, Genomic coordinates: Chromosomal coordinates of the identified binding sequence, Ensembl gene id, Id gene: Approved HUGO gene symbol for each protein coding gene, Tag location: Location of sequence within the gene, UTR type: Describes if the CLIP tag is present in an UTR, Splicing event: Alternative splicing annotation based upon Fast-DB database.

CLIP id	Genomic coordinates	Ensembl gene id	id gene	Tag location	UTR type	Splicing Event
517	chr6:170,697,141-170,697,181	ENSG00000008018	PSMB1	Exon 3		
669	chr3:56,572,949-56,572,968	ENSG00000180376	CCDC66	Exon 5	3'	Exon 4 skipping
672	chr3:53,504,274-53,504,295	ENSG00000157388	CACNAID	Exon 1		
682	chr2:206,735,019-206,735,052	ENSG00000114942	EEF1B2	Exon 5		
689	chr6:3,091,821-3,091,841	ENSG00000137274	BPHL	Intron 7		Alt 3'ss exon 7
707	chr21:31,990,834-31,990,853	ENSG00000156304	SFRS15	Exon 8		Alt 3'ss exon 8
726	chr17:45,401,725-45,401,746	ENSG00000108813	DLX4	Exon 1	5'	
738	chr19: 63,765,382-63,765,404	ENSG00000099326	MZF1	Exon 9		Alt 3'ss exon 9
755	chr5:6,721,930-6,721,954	ENSG00000145545	SRD5A1	Exon 5	5'	
826	chr5:6,721,930-6,721,954	ENSG00000145545	SRD5A1	Exon 5	5'	
832	chr7:92,300,701 - 92,300,720	ENSG00000105810	CDK6	Exon 2	5' and ORF	
848	chr14:22,601,486-22,601,513	ENSG00000100813	ACIN1	Exon 18		
851	chr14:22,602,590-22,602,624	ENSG00000100813	ACIN1	Exon 15		
234	chr14:22,598,412-22,598,445	ENSG00000100813	ACIN1	Exon 21		
803	chr14:22,608,570-22,608,605	ENSG00000100813	ACIN1	Exon11		
588	chr14:22,603,447-22,603,500	ENSG00000100813	ACIN1	Exon 13		
516	chr14:22,601,253-22,601,319	ENSG00000100813	ACIN1	Exon 19		
295	chr14:22,598,412-22,598,445	ENSG00000100813	ACIN1	Exon 21		
865	chr11:47,745,377-47,745,398	ENSG00000109920	FNBP4	Exon 1		Alt 5'ss exon 1
878	chr12:55,322,589-55,322,620	ENSG00000110955	ATP5B	Exon 7		
897	chr4:26,035,067-26,035,104	ENSG00000168214	RBPJ	Exon 10		
11	chrX:48,644,588-48,645,311	ENSG00000102103	PQBP1	Exon 8	3'	

Appendix 1 (continued).

CLIP id	Fraction	Genomic coordinates	Ensembl gene id	id gene	Tag location	UTR type	Splicing event
O11	Nuc -OSM	chr20:14,127,192-14,127,258	ENSG00000172264	MACROD2	Intron 3		
I11	Nuc -OSM	chr20:14,127,192-14,127,258	ENSG00000172264	MACROD2	Intron 3		
I17	Nuc -OSM	chr20:14,127,192-14,127,258	ENSG00000172264	MACROD2	Intron 3		
K09	Nuc -OSM	chr20:14,127,192-14,127,258	ENSG00000172264	MACROD2	Intron 3		
D08	Nuc -OSM	chr1:10,281,634-10,281,697	ENSG00000054523	KIF1B	Intron 24		Exon 25 skipping
G21	Nuc -OSM	chr1:10,281,634-10,281,697	ENSG00000054523	KIF1B	Intron 24		
A11	Nuc -OSM	chr10:28,864,495-28,864,563	ENSG00000095787	WAC	Exon 4		
4c10_2	Nuc -OSM	chr22:19,636,527-19,636,561	ENSG00000099942	CRKL	Intron 3		Intron 3 retention
B02	Nuc -OSM	chr12:54,723,146-54,723,202	ENSG00000197728	RPS26	Intron 2		
C03	Nuc -OSM	chr6:42,693,618-42,693,680	ENSG00000024048	UBR2	Intron 12		Exons 11 and 12 skipping
C08	Nuc -OSM	chr4:30,479,822-30,479,888	ENSG00000169851	PCDH7	Intron 2		
C11	Nuc -OSM	chr7:65,798,005-65,798,072	ENSG00000154710	KCTD7	Intron 3		
E17	Nuc -OSM	chr1:112,994,722-112,994,787	ENSG00000116489	CAPZA1	Intron 3		
E23	Nuc -OSM	chr9:130,390,973-130,391,032	ENSG00000197694	SPTAN1	Exon 21		Exons 16-56 skipping
O01	Nuc -OSM	chr1:82,229,287-82,229,348	ENSG00000117114	LPHN2	Exon 23		Exons 20-22 skipping
A08	Nuc -OSM	chr2:132,201,429-132,201,493	ENSG00000197927	C2ORF27	Intron 1		
A07	Nuc -OSM	chr1:24,166,174-24,166,222	ENSG00000215699	FUSIP1	Exon 6		
B05	Nuc -OSM	chr17:39,255,820-39,255,885	ENSG00000161647	MPP3	Intron 10		
B10	Nuc -OSM	chr12:3,191,190-3,191,260	ENSG00000130041	RPS27	Exon 1		
C04	Nuc -OSM	chr15:79,032,541-79,032,601	ENSG00000117899	MESDC2	Intron 3		
C10	Nuc -OSM	chr3:42,631,928-42,631,993	ENSG00000114857	NKTR	Intron 2		
D03	Nuc -OSM	chr11:6,480,476-6,480,548	ENSG00000179532	C11ORF47	Exon 4		
D10	Nuc -OSM	chr17:40,318,181-40,319,496	ENSG00000108883	EFTUD2	Exon 3		Internal exon 3 deletion
E03	Nuc -OSM	chrX:13,935,841-13,935,894	ENSG00000046647	GEMIN8	Exon 5	3'	
E08	Nuc -OSM	chr16:21,972,325-21,972,389	OTTHUMG00000131587	C16ORF52	Intron 1		
E09	Nuc -OSM	chr7:158,248,593-158,250,168	ENSG00000117868	FAM62B	Exons 11 and 12		
E12	Nuc -OSM	chr22:42,552,071-42,552,195	ENSG00000130540	SULT4A1	Exon 8	3'	Exon 7 skipping

Appendix 2. A list of annotation of the CLIP tags identified for hnRNP A1. Ensembl, UCSC Genome Browser and Fast-DB databases were used. Headers for the table are as follows: CLIP id: Id given to each tag during analysis, Fraction: Cellular fraction (nuclear or cytoplasmic) that the tag was identified from and whether this was from cells which had been osmotically shocked (OSM), Genomic coordinates: Chromosomal coordinates of the identified binding sequence, Ensembl gene id, Id gene: Approved HUGO gene symbol for each protein coding gene, Tag location: Location of sequence within the gene, UTR type: Describes if the CLIP tag is present in an UTR, Splicing event: Alternative splicing annotation based upon Fast-DB database.

CLIP id	Fraction	Genomic coordinates	Ensembl gene id	id gene	Tag location	UTR type	Splicing event
F01	Nuc -OSM	chr17:70,656,312-70,662,052	ENSG00000189159	HN1	Exons 1 and 4		Exons 2-4 skipping
G23	Nuc -OSM	chr2:86,966,978-86,967,046	ENSG00000153561	RMND5A	Intron 5		Exons 3-19 skipping
I01	Nuc -OSM	chr22:40,416,554-40,416,620	ENSG00000184208	FLJ23584	Exon 1		
M01	Nuc -OSM	chr1:160,444,148-160,444,228	ENSG00000198929	NOS1AP	Intron 2		
M05	Nuc -OSM	chr5:108,148,839-108,148,903	ENSG00000151422	FER	Intron 4		Exon 4 skipping
M09	Nuc -OSM	chr5:108,148,839-108,148,903	ENSG00000151422	FER	Intron 4		Exon 4 skipping
O15	Nuc -OSM	chr5:108,148,839-108,148,903	ENSG00000151422	FER	Intron 4		Exon 4 skipping
D04	Nuc -OSM	chr5:108,148,839-108,148,903	ENSG00000151422	FER	Intron 4		Exon 4 skipping
M17	Nuc -OSM	chr15:66,859,378-66,859,445	ENSG00000140350	ANP32A	Exon 7		
4BIS9	Nuc -OSM	chr15:66,877,021-66,877,055	ENSG00000140350	ANP32A	Intron 1		
4BIS8	Nuc -OSM	chr17:5,128,103-5,128,141	ENSG00000029725	RABEP1	Intron 1		
4C4.3	Nuc -OSM	chrX:118,440,611-118,440,646	ENSG00000077713	SLC25A43	Intron 3		Exon 4 skipping
4C5.2	Nuc -OSM	chr2:36,444,989-36,445,165	ENSG00000150938	CRIM1	Intron 1		
4C10.1	Nuc -OSM	chr1:152,681,318-152,681,349	ENSG00000160712	IL6R	Intron 7		Exon 7 skipping
451B	Nuc -OSM	chr16:28,122,802-28,122,841	ENSG00000169180	XPO6	Intron 2		Exon 2 skipping
489B	Nuc -OSM	chrX:10,452,530-10,452,560	ENSG00000101871	MID1	Intron 8		Exons 7-8 skippping
4C30.1	Nuc +OSM	chr2:132,228,880-132,228,899	ENSG00000197927	C2ORF27	Intron 1		
4C30.2	Nuc +OSM	chr7:106,665,571-106,665,605	ENSG00000164597	COG5	Intron 17		Exon 16 skipping
4C32.1	Nuc +OSM	chr16:9,928,119-9,928,138	ENSG00000183454	GRIN2A	Intron 4		
4C35	Nuc +OSM	chr6:20,729,933-20,729,955	ENSG00000145996	CDKAL1	Intron 5		Exon 5 skipping
460B	Nuc +OSM	chr17:73,943,215-73,943,284	ENSG00000187775	DNAH17	Intron 19		
K02	Cyt -OSM	chr6:146,967,064-146,967,116	ENSG00000118492	C6ORF103	Intron 1		
K10	Cyt -OSM	chr1:10,281,633-10,281,691	ENSG00000054523	KIF1B	Intron 24		
A06	Cyt -OSM	chr7:25,270,962-25,270,984	ENSG00000202233	SNORD3B-2	Exon 1		
O20	Cyt -OSM	chr2:85,624,567-85,624,629	ENSG00000168906	MAT2A	Exon 9	3'	
A06	Cyt -OSM	chr4:39,557,774-39,557,839	ENSG00000121892	PDS5A	Exon 19		
A14	Cyt -OSM	chr5:108,148,839-108,148,903	ENSG00000151422	FER	Intron 4		Exon 4 skipping
I10	Cyt -OSM	chr12:56,431,330-56,431,560	ENSG00000135446	CDK4	Exon 3		Exon 4 skipping; alt 5' ss exon 3
K22	Cyt -OSM	chr18:43,676,980-43,677,042	ENSG00000175387	SMAD2	Exon 3		
M10	Cyt -OSM	chr15:66,858,963-66,859,021	ENSG00000140350	ANP32A	Exon 7		

Appendix 2 (continued).

CLIP id	Fraction	Genomic coordinates	Ensembl gene id	id gene	Tag location	UTR type	Splicing event
M22	Cyt -OSM	chr7:150,552,108-150,552,838	ENSG00000033050	ABCF2	Exons 3 and 4		
O02	Cyt -OSM	chr7:129,465,528-129,465,601	ENSG00000091732	ZC3HC1	Intron 5		
A12	Cyt -OSM	chr17:30,502,348-30,502,388	ENSG00000141161	UNC45B	Intron 4		Exons 3-19 skipping
I08	Cyt -OSM	chr2:219,252,182-219,252,542	ENSG00000163482	STK36	Exons 7 and 8		
K16	Cyt -OSM	chr1:146,848,423-146,848,479	ENSG00000122497	NBPF14	Exon 3		
14.1	Cyt -OSM	chr10:64,681,532-64,681,564	ENSG00000171988	JMJD1C	Intron 4		Exon 3 skipping
24.2	Cyt -OSM	chr5:1,549,009-1,549,046	ENSG00000153395	AYTL2	Intron 2		
G2	Cyt +OSM	chr22:27,781,732-27,781,765	ENSG00000183579	ZNRF3	Exon 10	3'	
G37c	Cyt +OSM	chr22:27,781,732-27,781,765	ENSG00000183579	ZNRF3	Exon 10	3'	
G73c	Cyt +OSM	chr22:27,781,732-27,781,765	ENSG00000183579	ZNRF3	Exon 10	3'	
G86	Cyt +OSM	chr22:27,781,732-27,781,765	ENSG00000183579	ZNRF3	Exon 10	3'	
G4	Cyt +OSM	chr9:99,800,708-99,800,754	ENSG00000136938	ANP32B	Exon 3		
G40a	Cyt +OSM	chr9:99,800,708-99,800,754	ENSG00000136938	ANP32B	Exon 3		
G88	Cyt +OSM	chr9:99,800,708-99,800,754	ENSG00000136938	ANP32B	Exon 3		
G8	Cyt +OSM	chr5:74,686,104-74,686,155	ENSG00000113161	HMGCR	Exon 12		Exon 13 skipping
G53c	Cyt +OSM	chr5:74,686,104-74,686,155	ENSG00000113161	HMGCR	Exon 12		Exon 13 skipping
G92	Cyt +OSM	chr5:74,686,104-74,686,155	ENSG00000113161	HMGCR	Exon 12		Exon 13 skipping
G9	Cyt +OSM	chr12:52,962,448-52,962,480	ENSG00000135486	HNRNPA1	Exon 5		
G18	Cyt +OSM	chr12:52,962,448-52,962,480	ENSG00000135486	HNRNPA1	Exon 5		
G54a	Cyt +OSM	chr12:52,962,448-52,962,480	ENSG00000135486	HNRNPA1	Exon 5		
G62d	Cyt +OSM	chr12:52,962,448-52,962,480	ENSG00000135486	HNRNPA1	Exon 5		
G93	Cyt +OSM	chr12:52,962,448-52,962,480	ENSG00000135486	HNRNPA1	Exon 5		
G102	Cyt +OSM	chr12:52,962,448-52,962,480	ENSG00000135486	HNRNPA1	Exon 5		
G71a	Cyt +OSM	chr12:52,962,640-52,962,689	ENSG00000135486	HNRNPA1	Exon 6		Alt 5'ss exon 6; Exons 7-8 skipping
G61b	Cyt +OSM	chr12:52,962,448-52,962,480	ENSG00000135486	HNRNPA1	Exon 5		
G15	Cyt +OSM	chr12:52,962,448-52,962,480	ENSG00000135486	HNRNPA1	Exon 5		
G99	Cyt +OSM	chr12:52,962,448-52,962,480	ENSG00000135486	HNRNPA1	Exon 5		
G14	Cyt +OSM	chr21:37,807,225-37,807,274	ENSG00000157540	DYRK1A	Exon 14	3'	
G60b	Cyt +OSM	chr21:37,807,225-37,807,274	ENSG00000157540	DYRK1A	Exon 14	3'	
G98	Cyt +OSM	chr21:37,807,225-37,807,274	ENSG00000157540	DYRK1A	Exon 14	3'	

Appendix 2 (continued).

CLIP id	Fraction	Genomic coordinates	Ensembl gene id	id gene	Tag location	UTR type	Splicing event
G16	Cyt +OSM	chr4:77,879,164-77,879,208	ENSG00000138771	SHROOM3	Exon 7		Alt 5'ss exon 7; Alt 3'ss exon 7; exons 6 and 8 skipping
G100	Cyt +OSM	chr4:77,879,164-77,879,208	ENSG00000138771	SHROOM3	Exon 7		Alt 5'ss exon 7; Alt 3'ss exon 7; exons 6 and 8 skipping
G62b	Cyt +OSM	chr4:77,879,164-77,879,208	ENSG00000138771	SHROOM3	Exon 7		Alt 5'ss exon 7; Alt 3'ss exon 7; exons 6 and 8 skipping
G22	Cyt +OSM	chr14:57,908,849-57,908,888	ENSG00000032219	ARID4A	Exon 24	3'	
G68a	Cyt +OSM	chr14:57,908,849-57,908,888	ENSG00000032219	ARID4A	Exon 24	3'	
G106	Cyt +OSM	chr14:57,908,849-57,908,888	ENSG00000032219	ARID4A	Exon 24	3'	
G27	Cyt +OSM	chr17:7,756,143-7,756,197	ENSG00000170004	CHD3	Exon 40	3'	
G79a	Cyt +OSM	chr17:7,756,143-7,756,197	ENSG00000170004	CHD3	Exon 40	3'	
G111	Cyt +OSM	chr17:7,756,143-7,756,197	ENSG00000170004	CHD3	Exon 40	3'	
G37b	Cyt +OSM	chr13:97,619,508-97,619,557	ENSG00000152767	FARP1	Intron 2		
G73b	Cyt +OSM	chr13:97,619,508-97,619,557	ENSG00000152767	FARP1	Intron 2		
G40b	Cyt +OSM	chr1:235,833,114-235,833,142	ENSG00000198626	RYR2	Intron 33		
G45b	Cyt +OSM	chr1:235,833,114-235,833,142	ENSG00000198626	RYR2	Intron 33		
G70b	Cyt +OSM	chr1:235,833,114-235,833,142	ENSG00000198626	RYR2	Intron 33		
G51	Cyt +OSM	chr6:36,673,653-36,673,697	ENSG00000112081	SFRS3	Intron 2		
G54b	Cyt +OSM	chr17:55,270,705-55,270,752	ENSG00000062716	TMEM49	Intron 11		
G64b	Cyt +OSM	chr2:143,799,027-143,799,063	ENSG00000075884	ARHGAP15	Intron 6		
G66c	Cyt +OSM	chr8:56,956,958-56,956,995	ENSG00000147507	LYN	Intron 1		
G72b	Cyt +OSM	chr2:61,109,523-61,109,573	ENSG00000162928	PEX13	Intron 1		
4c40_1	Cyt +OSM	chr17:76,172,098-76,172,135	ENSG00000141564	RAPTOR	Intron 1		
G1	Cyt +OSM	chr1:55,370,902-55,370,938	ENSG00000162402	USP24	Exon 14		
G3	Cyt +OSM	chr16:87,313,857-87,314,006	ENSG00000103335	FAM38A	Exons 31 and 32		
G12	Cyt +OSM	chr11:5,203,432-5,203,483	ENST00000330597	HBG1	Exon 3		
G96	Cyt +OSM	chr11:5,203,432-5,203,483	ENST00000330597	HBG1	Exon 3		
G13	Cyt +OSM	chr15:98,969,934-98,969,983	ENSG00000183475	ASB7	Exon 4		
G97	Cyt +OSM	chr15:98,969,934-98,969,983	ENSG00000183475	ASB7	Exon 4		
G17	Cyt +OSM	chr17:30,481,032-30,481,097	ENSG00000073598	FNDC8	Intron 3		
G101	Cyt +OSM	chr17:30,481,032-30,481,097	ENSG00000073598	FINDC8	Intron 3		

Appendix 2 (continued).

CLIP id	Fraction	Genomic coordinates	Ensembl gene id	id gene	Tag location	UTR type	Splicing event
G24	Cyt +OSM	chr19:44,618,149-44,618,190	ENSG00000105193	RPS16	Exon 2		Alt 5'ss exon 2
G108	Cyt +OSM	chr19:44,618,149-44,618,190	ENSG00000105193	RPS16	Exon 2		Alt 5'ss exon 2
G91	Cyt +OSM	chr11:62,095,672-62,095,705	ENSG00000186676	EEF1G	Exon 4		
G95	Cyt +OSM	chr1:152,199,138-152,199,164	ENSG00000143570	SLC39A1	Exon 4	3'	
G105	Cyt +OSM	chr20:48,175,087-48,175,126	ENSG00000124208	TMEM189/UBE2V1	Exon 6		Exons 5-10 skipping
G110	Cyt +OSM	chr12:10,266,219-10,266,250	ENSG00000139112	GABARAPL1	Exon 7	3'	
G112	Cyt +OSM	chr5:17,329,585-17,329,620	ENSG00000176788	BASP1	Exon 2	3'	Internal exon 2 deletion
G42b	Cyt +OSM	chr11:65,029,271-65,029,310		NEAT2			
G45a	Cyt +OSM	chr11:65,026,992-65,027,036		NEAT2			
G53b	Cyt +OSM	chr11:65,027,628-65,027,671		NEAT2			
D07	Nuc -OSM	chr1:16,939,626-16,939,697	ENSG00000207389	U1A snRNA	Exon 1		
B01	Nuc -OSM	chr1:16,939,626-16,939,697	ENSG00000207389	U1A snRNA	Exon 1		
I23	Nuc -OSM	chr1:16,939,626-16,939,697	ENSG00000207389	U1A snRNA	Exon 1		
G21	Nuc -OSM	chr1:10,281,633-10,281,691	ENSG00000199562	U6snRNA	Exon 1		
D08	Nuc -OSM	chr1:10,281,634-10,281,697	ENSG00000199562	U6 snRNA	Exon 1		
G04	Cyt -OSM	chr22: 41,341,242- 41,341,307		U12snRNA	Exon 1		
K23	Nuc -OSM	chr22:41,341,253-41,341,307		U12snRNA	Exon 1		
469A	Nuc -OSM	chr13:90,801,034-90,801,076	ENSG00000215417	MIRH1	Exon 3		Intron 3 retention
4C31.2	Nuc +OSM	chr13:90,801,041-90,801,076	ENSG00000215417	MIRH1	Exon 3		Intron 3 retention
4C39.1	Nuc +OSM	chr13:90,801,040-90,801,076	ENSG00000215417	MIRH1	Exon 3		Intron 3 retention
4105A	Nuc +OSM	chr13:90,801,040-90,801,076	ENSG00000215417	MIRH1	Exon 3		Intron 3 retention
4bis14_2	Cyt -OSM	chr13:90,801,039-90,801,076	ENSG00000215417	MIRH1	Exon 3		Intron 3 retention
468	Cyt +OSM	chr13:90,801,034-90,801,076	ENSG00000215417	MIRH1	Exon 3		Intron 3 retention
G10	Cyt +OSM	chr6:52,968,641-52,968,676	ENSG00000202198	RN7SK	Exon 1		
G54c	Cyt +OSM	chr6:52,968,641-52,968,676	ENSG00000202198	RN7SK	Exon 1		
G94	Cyt +OSM	chr6:52,968,641-52,968,676	ENSG00000202198	RN7SK	Exon 1		
4c12_1	Nuc -OSM	chr6:52,968,637-52,968,671	ENSG00000202198	RN7SK	Exon 1		
G24	Cyt -OSM	chrMT:11,958-12,021	ENSG00000198886	MT-ND4	Exon 1		
I20	Cyt -OSM	chrMT:8,945-9,009	ENSG00000198899	MT-ATP6	Exon 1		

Appendix 2 (continued).

CLIP id	Fraction	Genomic coordinates	Ensembl gene id	id gene	Tag location	UTR type	Splicing event
G23	Cyt +OSM	chr1:91,920,296-91,920,349	intergenic				
4c48.2	Cyt +OSM	chr16:88,249,038-88,249,057	intergenic				
4c46	Cyt +OSM	chr7:64,532,328-64,532,422	intergenic				
A02	Cyt -OSM	chr14:70,912,270-70,912,328	intergenic				
G10	Cyt -OSM	chr5:173803240-173803307	intergenic				
C09	Nuc -OSM	chr3:4,366,009-4,366,068	intergenic				
M11	Nuc -OSM	chr11:62,116,816-62,116,874	intergenic		Tag 5' of TUT1		
M23	Nuc -OSM	chr17:53,440,241-53,440,295	intergenic		Tag 5' of SFRS1		
494A	Nuc -OSM	chrX:10,809,239-10,809,275	intergenic				
495A	Nuc -OSM	chr5:56,482,989-56,483,016	intergenic				
4C29.2	Nuc +OSM	chr9:129,808,183-129,808,219	intergenic				

Appendix 2 (continued).