Regulation of In Vitro Pre-mRNA Splicing by hnRNP A1 Cooperative Binding

A Dissertation Presented
by
Hazeem L Okunola

to
The Graduate School
in Partial Fulfillment of the
Requirements
of the Degree of
Doctor of Philosophy
in
Physiology and Biophysics
Stony Brook University

December 2008
We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of this dissertation.

Dr. Adrian R Krainer – Dissertation Advisor
Professor – Cold Spring Harbor Lab

Dr. Raafat El-Maghrabi – Chairperson of Defense
Associate Professor – Department of Physiology and Biophysics

Dr. Nicolas Nassar – Committee Member
Associate Professor – Department of Physiology and Biophysics

Dr. Peter Brink – Committee Member
Professor and Chairman – Department of Physiology and Biophysics

Dr. Wali Karzai – Committee Member
Associate Professor – Department of Biochemistry and Cell Biology

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School
Abstract of the Dissertation

Regulation of in Vitro Pre-mRNA Splicing by hnRNP A1 Cooperative Binding

by

Hazeem L Okunola

Doctor of Philosophy

in

Physiology and Biophysics

Stony Brook University

2008

Splicing is defined as the removal of introns and joining together of exons from primary transcripts. Splicing, which is essential for eukaryotic gene expression, can be constitutive or alternative. In constitutive splicing, all the exons are joined together in the same order in which they are present along a gene. In alternative splicing, part of an exon or the whole exon can be included or skipped in the final, mature mRNA products. Alternative splicing is the principal means by which eukaryotes diversify the number of proteins expressed from a single gene. Many alternative splicing events are regulated by the interplay between ESE (exonic splicing enhancer) elements that bind SR proteins (Serine/Arginine-rich proteins) and ESS (exonic splicing silencer) elements that bind hnRNPs (heterogeneous nuclear ribonucleoproteins), notably hnRNP A1. hnRNP A1 inhibits splicing in a cooperative-binding-dependent manner. When hnRNP A1 is able to displace bound SR protein from an ESE, the result is skipping of the corresponding exon. In contrast, when the SR protein binds tightly enough to disrupt hnRNP A1 cooperative binding along the exon, the exon is included. I show in Chapter 2 that hnRNP A1 cooperative binding does not require RNA secondary structure, and that this cooperative binding along an RNA can displace a protein tightly bound to an RNA hairpin, and unwind the hairpin structure. I also show that hnRNP A1 cooperative binding can spread along RNA from 5’ to 3’, in addition to the known 3’ to 5’ spreading. This type of binding, similar to beads on a string, results in repression of splicing. When hnRNP A1 initial binding takes place in the middle of an RNA, spreading proceeds
preferentially in a 3’ to 5’ direction. I also show that there is an interaction, or cross-talk, between hnRNP A1 bound at two distant sites, through cooperative binding. In Chapter three, I describe methods and experiments to derive a more accurate hnRNP A1 consensus motif for sequence-specific, high-affinity binding.
TABLE OF CONTENTS

List of Figures ........................................................................................................... vii
List of Abbreviations ............................................................................................... viii
Acknowledgments .................................................................................................... x

Chapter 1 Background ............................................................................................... 1

1.1 Pre-mRNA Splicing .......................................................................................... 2
  1.1.1 Introduction to Pre-mRNA Splicing ......................................................... 2
  1.1.2 Constitutive Splicing ............................................................................. 2
  1.1.3 Alternative Splicing ............................................................................. 3
  1.1.4 DNA Transcription and Splicing ........................................................... 5
  1.1.5 Mature mRNA Export ........................................................................ 5

1.2 Splicing Machinery ......................................................................................... 6
  1.2.1 Splice Sites ........................................................................................... 6
  1.2.2 Spliceosome Assembly ...................................................................... 6

1.3 Regulation of Splicing ................................................................................... 9
  1.3.1 Cis-acting Elements .......................................................................... 9
    A. Exonic Splicing Enhancers .................................................................. 9
    B. Exonic Splicing Silencers .................................................................. 9
  1.3.2 Trans-acting Factors ........................................................................ 10
    A. SR proteins ......................................................................................... 10
    B. hnRNP proteins ................................................................................. 14

1.4 Antagonism between Positive and Negative Regulators ......................... 15

1.5 References ....................................................................................................... 15

Chapter 2 Mechanism of hnRNP A1 Cooperative Binding .............................. 28

2.1 Abstract .......................................................................................................... 29

2.2 Introduction ..................................................................................................... 29

2.3 Results ............................................................................................................. 32
  2.3.1 Cooperative binding of hnRNP A1 does not require RNA secondary structure ................................................................................................................. 32
  2.3.2 hnRNP A1 unwinds secondary structure ............................................ 35
  2.3.3 Cooperative binding of hnRNP A1 can also proceed from 5' to 3' and inhibits splicing ................................................................. 37
  2.3.4 Determining the extent of spreading of hnRNP A1 along the RNA ......................................................................................................................... 39
2.3.5 hnRNP A1 preferentially spreads in a 3' to 5' direction ................. 40
2.3.6 "Cross-talk" between hnRNP A1 molecules bound at distant sites ............................................................................................................. 43

2.4 Discussion ................................................................................................. 45

2.5 Methods and Materials ............................................................................. 49
2.5.1 Transcripts ............................................................................................... 49
2.5.2 Recombinant Proteins ............................................................................. 51
2.5.3 In Vitro Splicing Assays ........................................................................ 51
2.5.4 UV Crosslinking ..................................................................................... 52
2.5.5 GST-MS2 Pulldown ............................................................................... 52
2.5.6 RNA Footprinting .................................................................................. 53

2.6 Acknowledgments ...................................................................................... 53

2.7 References .................................................................................................. 54

Chapter 3 Consensus Sequences of hnRNP A1 ............................................ 58
3.1 Abstract ...................................................................................................... 59

3.2 Introduction .................................................................................................. 60

3.3 Results ......................................................................................................... 61
3.3.1 Nβ-globin Constructs .......................................................................... 61
3.3.2 Splicing Assay Results ........................................................................ 62
3.3.3 Matrix and the Consensus Sequences for hnRNP A1 ....................... 71

3.4 Discussion .................................................................................................... 72

3.5 Methods and Materials ............................................................................. 73
3.5.1 Transcripts ............................................................................................... 73
3.5.2 Recombinant Proteins ............................................................................. 73
3.5.3 In Vitro Splicing Assays ........................................................................ 73

3.6 References .................................................................................................... 74

Chapter 4 Summary ......................................................................................... 77
4.1 Summary ...................................................................................................... 78

4.2 References .................................................................................................... 79

Bibliography ...................................................................................................... 80
List of Figures

Figure 1. Diagram of constitutive and alternative pre-mRNA splicing.
Figure 2. Spliceosome assembly.
Figure 3. SR proteins.
Figure 4. Structure of hnRNP A1 showing the domains.
Figure 5. Cooperative binding of hnRNP A1 does not require RNA secondary structure.
Figure 6. Unwinding of RNA secondary structure and splicing inhibition by hnRNP A1.
Figure 7. hnRNP A1 cooperative binding spreading from 5’ end to 3’ end of exon 1 inhibit splicing.
Figure 8. Mapping the extent of hnRNP A1 cooperative spreading along the RNA.
Figure 9. Directionality of hnRNP A1 cooperative spreading.
Figure 10. “Cross-talk” between hnRNP A1 molecules bound at distant sites.
Figure 11. Model of hnRNP A1 cooperative binding.
Figure 12. Nβ-globin constructs.
Figure 13. Splicing assay results. Pool 1
Figure 14. Splicing assay results. Pool 1 Control
Figure 15. Splicing assay results. Pool 2
Figure 16. Splicing assay results. Pool 2 Control
Figure 17. Splicing assay results. Pool 3
Figure 18. Splicing assay results. Pool 3 Control
Figure 19. Splicing assay results. Pool 4
Figure 20. Splicing assay results. Pool 4 Control
Figure 21. Matrix representing hnRNP A1 binding consensus sequences.
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBP</td>
<td>branch-point binding protein</td>
</tr>
<tr>
<td>BPS</td>
<td>branch-point sequence</td>
</tr>
<tr>
<td>CBC</td>
<td>cap-binding complex</td>
</tr>
<tr>
<td>CTD</td>
<td>C-terminal domain</td>
</tr>
<tr>
<td><em>dsx</em></td>
<td>double-sex gene</td>
</tr>
<tr>
<td>ESE</td>
<td>exonic splicing enhancer</td>
</tr>
<tr>
<td>ESS</td>
<td>exonic splicing silencer</td>
</tr>
<tr>
<td>EST</td>
<td>expressed sequence tag</td>
</tr>
<tr>
<td>FGF</td>
<td>Fibroblast growth factor</td>
</tr>
<tr>
<td>hnRNP</td>
<td>heterogeneous nuclear ribonucleoprotein</td>
</tr>
<tr>
<td>IgM</td>
<td>Immunoglobulin µ</td>
</tr>
<tr>
<td>ISE</td>
<td>intronic splicing enhancer</td>
</tr>
<tr>
<td>ISS</td>
<td>intronic splicing silencer</td>
</tr>
<tr>
<td>IVS</td>
<td>intervening sequence</td>
</tr>
<tr>
<td>Luc</td>
<td>Lethal unless CBC is produce</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>NMD</td>
<td>nonsense-mediated mRNA degradation</td>
</tr>
<tr>
<td>PAP</td>
<td>Poly(A) polymerase</td>
</tr>
<tr>
<td>PPT</td>
<td>polypyrimidine tract</td>
</tr>
<tr>
<td>Pre-mRNA</td>
<td>precursor messenger RNA</td>
</tr>
<tr>
<td>PTB</td>
<td>polypyrimidine tract binding protein</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RRM</td>
<td>RNA recognition motif</td>
</tr>
<tr>
<td>RBD</td>
<td>RNA binding domain</td>
</tr>
<tr>
<td>RS domain</td>
<td>arginine/serine-rich domain</td>
</tr>
<tr>
<td>SCAF</td>
<td>SR-like CTD-associated protein</td>
</tr>
<tr>
<td>SELEX</td>
<td>systematic evolution of ligands by exponential enrichment</td>
</tr>
<tr>
<td>SF1</td>
<td>Splicing factor 1</td>
</tr>
<tr>
<td>Sip</td>
<td>SC35 interacting protein</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>snRNP</td>
<td>small nuclear ribonucleoprotein</td>
</tr>
<tr>
<td>SRPK</td>
<td>SR protein kinase</td>
</tr>
<tr>
<td>Sxl</td>
<td>sex lethal gene</td>
</tr>
<tr>
<td>tra</td>
<td>transformer</td>
</tr>
<tr>
<td>U2AF</td>
<td>U2 auxiliary factor</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>Urp</td>
<td>U2AF related protein</td>
</tr>
</tbody>
</table>
Acknowledgments

I am very thankful to my advisor, Dr. Adrian R. Krainer, for his guidance, patience and advice and all his helpful suggestions, especially during my first two years when nothing seemed to be working, and for always being there during the times of failed experiments and during the successful ones, to share my sorrows and joys.

I also thank all my lab mates for their contributions to my research, especially their contributions in the discussion and critical analysis of my data; therefore, I thank Jun Zhu, Stephanie Shaw, Yimin Hua, Philip Smith, Luca Cartegni, Michelle Hastings, Qingshuo Zhang, Zuo Zhang, Xavier Roca, Eric Allemand, Mads Jensen, Olga Anczukow-Camarda, Isabel Aznarez, Oliver Fregoso, Shuying Sun, Rahul Sinha, and Chaolin Zhang who helped me. My special thanks go to Lisa Manche, who always makes sure we have all the reagents and equipment necessary to do our work; my special phrase in the lab is “Thank God for Lisa” because we always have what we need in the lab.

I thank all the members of my thesis committee: my committee chairman, Dr. Raafat El-Maghrabi, Dr. Nicholas Nassar, Dr. Peter Brink, and Dr. A. Wali Karzai for all their helpful suggestions and guidance.

I thank all the members of the Department of Physiology and Biophysics for all their help over the years, and my special thanks go to Melanie Bonnette, who is always there to answer all my questions, and to Dr. Suzanne Scarlata, who gave me the opportunity to be one of the graduate students of the department.

Finally, I thank my parents for all their support.
Chapter 1: Background
1.1 Pre-mRNA Splicing

1.1.1 Introduction to pre-mRNA Splicing

Pre-mRNA splicing is a required process for proper gene expression in unicellular and multicellular eukaryotes. Pre-mRNA splicing involves the removal of intervening sequences, known as introns, and the joining of the coding sequences, known as exons, from a primary transcript to form a mature mRNA product. The introns that are spliced out are in the form of lariat-shaped RNA (Grabowski, Padgett & Sharp, 1984; Padgett et al., 1986; Ruskin et al., 1984). This process can be in the form of constitutive splicing, in which all the exons are joined in the same order in which they are present along a gene; or it can be in the form of alternative splicing, in which part of an exon or a whole exon is either skipped or included from the final mature mRNA product(s). Alternative splicing is a means by which eukaryotes diversify the number of protein expressed from a single gene. Many alternative splicing events are regulated by the interplay between ESE (exonic splicing enhancer) elements that bind SR (Serine/Arginine rich) proteins and ESS (exonic splicing silencer) elements that binds hnRNPs (heterogeneous nuclear ribonucleoproteins). If hnRNP binding prevails over SR protein binding and prevents splicing activation via the ESE, the result is skipping of the alternative exon; on the other hand, if the SR protein successfully prevents hnRNP-mediated repression, the result is inclusion of the alternative exon.

1.1.2 Constitutive Splicing

Constitutive splicing is the removal of the introns and joining together of the adjacent exons in a manner that ensures that no exon is skipped, as shown in Figure 1A. This type of splicing leads to the production of a single protein from a single gene; an example is the splicing of the three exons of the human beta-globin gene transcripts to generate a single beta-globin protein chain (Crick, 1979; Green, 1991; Orkin & Kazazian, 1984; Vidaud et al., 1989).
1.1.3 Alternative Splicing

Alternative splicing is the skipping of an exon or part of an exon during the process of removing intervening sequences from the primary transcript (Figure 1B). The mature messenger RNA (mRNA) that results from this splicing event is then transported to the cytoplasm for translation into protein. Since alternative splicing makes differential use of exon-intron junctions, many mRNAs may be produced from a single primary transcript. This is a way many organisms use to diversify the number of proteins produced from a single transcript (Green, 1991). An extreme example of this is the Drosophila Dscam gene, in which a single pre-mRNA transcript produces up to 38,016 protein isoforms through combinatorial alternative splicing events (Graveley, 2005; Kreahling & Graveley, 2005). Another example of alternative splicing is sex determination in Drosophila, in which a series of alternative splicing events affecting sex-specific transcription factors, allows sex differentiation of male and female Drosophila (Bell et al., 1991; Boggs et al., 1987; Hedley & Maniatis, 1991; Hoshijima et al., 1991; Ryner & Baker, 1991). Alternative splicing can in many cases be subjected to regulation, for example in a cell-type-specific manner, during embryonic development, or in response to signaling pathways.

Retroviruses like HIV-1 depend greatly on alternative splicing to produce all of the viral proteins from a single primary transcript (Caputi & Zahler, 2002). The unspliced transcript is necessary for viral replication, packaging into virions, and translation of several proteins, while other viral proteins are generated from partially spliced or fully spliced transcripts. Special mechanisms allow these incompletely spliced transcripts to be exported to the cytoplasm for translation.

About 50% of point mutations associated with human genetic diseases results in defective splicing of the mutant genes (Cartegni, Chew & Krainer, 2002; Cartegni et al., 1996). Analysis of the human genome indicated that more than 74% of human genes encode at least two isoforms by alternative splicing (Johnson et al., 2003; Kan, States & Gish, 2002; Lander et al., 2001; Okazaki et al., 2002).
A. Constitutive pre-mRNA splicing

B. Alternative splicing

**Figure 1.** Diagram of constitutive (A) and alternative (B) pre-mRNA splicing. The colored rectangular boxes represent pre-mRNA coding sequences known as exons, and the horizontal lines separating the exons are the intervening sequences, known as introns, that are removed during splicing. During this type of splicing event, all exons are joined together as all the introns are removed.
1.1.4 DNA Transcription and RNA Splicing

Within the nucleus are structures known as speckles, which reflect the organization of spliceosomes and other splicing factors. The splicing factors begin to move to the site of DNA transcription when RNA polymerase II transcription activation is detected (Misteli, Caceres & Spector, 1997). Movement of the splicing factors to the site of transcription activation shows the coupling of DNA transcription and RNA splicing (Jimenez-Garcia & Spector, 1993). The primary transcripts newly synthesized by RNA polymerase II are covalently modified at both the 5’ and the 3’ end. As the elongation of the newly transcribed RNA progresses to about 30 nucleotides, the 5’ end of the RNA is capped with a methylated guanine nucleotide. The 5’ cap of this newly synthesized RNA serves three purposes; the first is to differentiate the RNA newly transcribed by polymerase II from other RNAs transcribed by other polymerases; the second is to prevent this newly synthesized RNA from being degraded by nucleases; and the third is to serve as the initiation point for the protein translation machinery in the cytoplasm (Nevins, 1983; Proudfoot, 1989; Takagaki, Ryner & Manley, 1988; Wickens, 1990).

The 3’ end of the RNA newly synthesized by RNA polymerase II is polyadenylated by a poly-A-polymerase at a cleavage site located 10 – 30 nucleotides downstream of the polyadenylation signal AAUAAA. Up to 200 residues of non-templated adenylic acid (poly-A) may be added to the cleaved 3’ end of the RNA by poly-A-polymerase. This polyadenylation of the 3’ end of the primary transcript may serve as the recognition of mature mRNA for transport into the cytoplasm, and may also be recognized by the translation machinery, as well as help stabilize the mRNA in the cytoplasm (Jimenez-Garcia & Spector, 1993; Nevins, 1983; Proudfoot, 1989; Spector, 1993; Spector, Landon & O'Keefe, 1993a; Spector, O'Keefe & Jimenez-Garcia, 1993b; Takagaki et al., 1988; Wickens, 1990).

1.1.5 Mature mRNA Export

Mature mRNAs are exported into the cytoplasm at the completion of splicing, complexed with proteins, including hnRNPs. Many of these proteins are stripped from the mRNA (or exchanged for other proteins) upon reaching the cytoplasm, and most of them return to the nucleus
1.2 Splicing Machinery

1.2.1 Splice Sites

Intron splice sites are chosen or recognized in pairs: the 5' splice site and 3' splice site, which are characterized in part by highly conserved dinucleotides (GU and AG, respectively). The recognition of these splice sites depends in part on two types of snRNPs (small nuclear ribonucleoprotein particles), U2 and U12 snRNPs. The complex formed between small nuclear RNA (snRNA) and proteins makes up a snRNP (Roca, Sachidanandam & Krainer, 2003; Roca, Sachidanandam & Krainer, 2005; Sheth et al., 2006; Will & Luhrmann, 2001a; Will & Luhrmann, 2001b; Will et al., 2001). U2 snRNP-dependent introns have termini with dinucleotide configuration of GU – AG, and in some cases GC – AG, or AU – AC; U12 snRNP-dependent introns begin and end with AU – AC or GU – AG (Roca et al., 2003; Roca et al., 2005; Sheth et al., 2006). The most common and most conserved intronic pairs of splice sites are the ones recognized by U2 snRNP (Roca et al., 2003; Roca et al., 2005; Sheth et al., 2006). The 5' splice site, which is also known as the donor site, is recognized by the U1 or U11 snRNP through complementary base pairing with U1 or U11 snRNA (Roca et al., 2003; Roca et al., 2005; Sheth et al., 2006; Will & Luhrmann, 2001a; Will & Luhrmann, 2001b; Will et al., 2001).

1.2.2 Spliceosome assembly
The snRNPs associated with the splicing of U2-dependent introns are U1, U2, U4, U5, and U6 snRNPs, whereas the ones associated with splicing of U12-dependent introns are U11, U12, U4atac, U6atac, and U5 snRNPs (Sheth et al., 2006). For effective spliceosome assembly on a pre-RNA, at least four unique elements are required: the 5’ splice site (5’ss), a branch point sequence (BPS), followed by a polypyrimidine tract (PPT) and a 3’ splice site (3’ss) (Hastings & Krainer, 2001). Spliceosome assembly is organized in a stepwise manner with the formation of intermediate complexes called E, A, B and C (Figure 2). The E complex is formed when recognition of the 5’ss occurs through U1 snRNA base-pairing with the 5’ss. In addition, an SR protein that promotes binding of the U2 auxiliary factor (U2AF) to the PPT in an ATP-independent manner, and U2 snRNP might also be part of the E complex in a way that does not yet involve interaction with the BPS (Das, Zhou & Reed, 2000; Hastings & Krainer, 2001). Next is the A complex, in which U2 snRNA base pairing with the BPS occurs in an ATP-dependent manner. The interaction of the tri-snRNP, U4/U6·U5, with the pre-mRNA to form the B complex and the C complex occurs when the remodeling of RNA-RNA and protein-RNA of the B complex forms a catalytic spliceosome.

Even though U12-dependent intron splicing represents less than 1% of all splicing (Burge, Padgett & Sharp, 1998; König et al., 2007; Levine & Durbin, 2001), the organization of the splicing complexes may be similar to the U2-dependent intron splicing, except for the fact that U12 instead of U2 snRNA base pairs to the BPS and U11 instead of U1 snRNA base pairs to the 5’ss. Also, the formation of a catalytic core, U6atac and U12/U6atac, can occur without the displacement of U11 from the 5’ss by U6atac (Frilander & Steitz, 2001; Will et al., 2001).
Spliceosome assembly

Figure 2. Spliceosome assembly. The diagrams show the spliceosome complexes that assemble to form a catalytically competent spliceosome (reviewed by Hastings and Krainer, 2001).
1.3 Regulation of Splicing

1.3.1 Cis-acting elements

A. Exonic Splicing Enhancers

Exonic splicing enhancers (ESEs) are regulatory cis acting elements found within the exons (Blencowe, 2000), but the short nucleotide sequences that constitute an ESE are sometimes found within the intronic sequences as well. ESEs were first identified in the alternatively spliced *Drosophila* doublesex (*dsx*) gene as a purine-rich element in exon 4 that controls the inclusion of this female-specific exon (Lopez, 1998). ESEs were later identified in higher eukaryotes in numerous genes (Huh & Hynes, 1994; Lavigneuer et al., 1993; Lynch & Maniatis, 1995; Sun, Hampson & Rottman, 1993a; Sun et al., 1993b; Watakabe, Tanaka & Shimura, 1993). Purine-rich elements within the exon are likely areas to find ESEs, especially if they contain GA or GAA repeats (Dirksen et al., 1994; Watakabe et al., 1993). ESEs are known to bind the SR-family and SR-related proteins to promote both constitutive and alternative splicing (Blencowe, 2000; Graveley, 2000). Another function of ESEs is the enhancement of splicing of introns with weak splice sites, which otherwise may splice inefficiently or not splice at all, so the more ESEs present within an exon, the greater the enhancement of splicing of that exon (Hertel & Maniatis, 1998). The function of multisite enhancers may be to increase the probability of interaction between SR proteins and the splicing machinery, rather than to give increased functionality to the ESEs (Graveley, Hertel & Maniatis, 1999). In *Saccharomyces cerevisiae*, functional ESEs and SR proteins are absent and almost all the splice sites are very strong and conform mostly to the canonical splice-site consensus (Pleiss et al., 2007a; Pleiss et al., 2007b). Therefore, enhancement of splicing is not necessary.

B. Exonic Splicing Silencers

Exonic splicing silencers (ESSs) are regulatory cis-acting elements found within the exons, and lead to the exclusion of these exons in alternative splicing events. Most of the ESSs that have been identified (Amendt, Si & Stoltz fus, 1995; Caputi et al., 1994; Chew, Baginsky & Eperon, 2000; Del Gatto-Konczak et al., 1999; Zheng et al., 2000) bind
heterogeneous nuclear ribonucleoprotein family members (hnRNPs) (Amendt et al., 1995; Burd & Dreyfuss, 1994; Del Gatto-Konczak et al., 1999; Expert-Bezancon et al., 2004; Ma et al., 2002; Min, Chan & Black, 1995; Nasim et al., 2002; Paradis et al., 2007; Zhu, Mayeda & Krainer, 2001). The mechanism by which ESSs inhibit pre-mRNA splicing is still unclear, but a likely mechanism of inhibition is that the ESSs bind repressor proteins to form an inhibitory complex that prevents spliceosome assembly (Caputi et al., 1999; Chen, Kobayashi & Helfman, 1999; Del Gatto-Konczak et al., 1999). Most of the sequences with ESS activity can also be found within the introns of many genes as intronic splicing silencers (ISSs) which inhibit a particular exon. Examples include inhibition of exon 3 of α-tropomyosin (Gooding, Roberts & Smith, 1998), exon N1 of c-src (Markovtsoy et al., 2000; Min et al., 1995), exon IIIb of FGF-R2 (Carstens, McKeehan & Garcia-Blanco, 1998; Carstens, Wagner & Garcia-Blanco, 2000), and exon 7B of hnRNP A1 (Simard & Chabot, 2000). ISSs have also been found to inhibit the splicing of the nearby introns (Carstens et al., 2000; Chabot et al., 1997; Min et al., 1995; Simard & Chabot, 2000). Another ISS mechanism has been proposed, which involves the looping out of an exon between two introns with similar ISS that bind the same inhibitory protein (Nasim et al., 2002).

1.3.2 Trans-acting factors

A. SR proteins

Serine-arginine (SR) proteins are a family of trans-acting factors that usually bind to ESEs to activate splicing during constitutive and/or alternative splicing events. They recruit other splicing factors and co-activators of splicing onto pre-mRNA with weak splice sites (Blencowe, 2000). ESEs are not always necessary for splicing activation by SR proteins, such as in Saccharomyces cerevisiae (Pleiss et al., 2007a; Pleiss et al., 2007b), which lacks SR proteins. The first SR protein to be purified from HeLa cell nuclear extract was SF2/ASF, based on its ability to complement S100, a splicing-inactive cytoplasmic extract, in constitutive splicing, or to modulate alternative splicing in nuclear extract (Ge & Manley, 1990; Krainer, Conway & Kozak, 1990a; Krainer, Conway & Kozak, 1990b). Over the years, more SR protein family members have been identified (Roth, Zahler & Stolk, 1991; Screaton et al., 1995; Soret et al., 1998; Yang et al., 1998; Zahler et al., 1992; Zahler et al., 1993; Zhang & Wu, 1996) (Figure 4). All members of the SR protein family share two distinct domains: an N-terminal RNA-recognition motif (RRM) and a C-
terminal arginine-serine-rich (RS) domain (Birney, Kumar & Krainer, 1993). There is either one or two RRM s in all SR proteins (Figure 3).

The function of the RRM s is to determine the substrate specificity of the SR proteins (Chandler et al., 1997; Mayeda et al., 1999), and there are substrates that do not require the RS domain of the SR protein for splicing (Shaw et al., 2007; Zhu & Krainer, 2000), even though it is generally believed that the RS domain is involved in protein-protein and/or protein-RNA interactions. Protein-protein interaction between SR proteins was first shown in the Drosophila splicing regulators Tra and Tra2 and also in the interaction between the SR proteins and RS-like domains in other factors, such as U1-70K and U2AF35 (Kohtz et al., 1994; Wu & Maniatis, 1993). Also, both the RRM and RS domains are interchangeable with their homologous domains between SR proteins, indicating that they are modular domains (Chandler et al., 1997; Mayeda et al., 1999; Wang, Xiao & Manley, 1998b). The RS domain of an SR protein can activate splicing when fused with another protein’s RRM or RBD (RNA binding domain); for example, an RS domain fused with MS2 coat protein RBD will active splicing in a substrate containing an MS2 binding site in place of an ESE (Graveley & Maniatis, 1998); however, SR proteins are still required for this splicing reaction.

There are other proteins involved in splicing that are not canonical SR proteins, even though they possess an RS domain; these proteins are referred to as SR-related proteins (SRPs) (Blencow et al., 1999; Fu, 1995). Examples of these SRPs are U2AF (Zamore, Patton & Green, 1992; Zhang et al., 1992), snRNP U1-70K (Query, Bentley & Keene, 1989), snRNP U5-100K (Teigelkamp et al., 1997), snRNP U4/U6·U5-27K (Fetzer et al., 1997), hLuc7p (Fortes et al., 1999; Nishii et al., 2000), SWAP (Denhez & Lafyatis, 1994; Spikes et al., 1994), SRm160/SRm300 (Blencow et al., 2000; Blencow et al., 1998), the RNA helicases hPrp16 (Zhou & Reed, 1998), HRH1 (Ono, Ohno & Shimura, 1994), Hel 117 (Sukegawa & Blobel, 1995), and also the protein kinases Clk/Sty1, 2, 3 (Hanes et al., 1994), and the splicing regulators Tra (Boggs et al., 1987), Tra2 (Amrein, Gorman & Nothiger, 1988), hTra2α and hTra2β (Beil, Screaton & Stamm, 1997).

SR proteins can be post-translationally modified, especially on their arginine and serine residues. Arginine residues can be methylated and dimethylated (Ong, Mittler & Mann, 2004) (Sinha and Krainer, unpublished) and serine residues may be phosphorylated. Some of the kinases that phosphorylate SR proteins have been identified, and include the SR protein kinases SRPK1 (Gui et al., 1994) and SRPK2 (Wang et al., 1998a), Clk/Sty (Colwill et al., 1996), and DNA topoisomerase I (Rossi et al., 1996). Dephosphorylation of an SR protein, SRp40, causes it to lose its RNA binding activity (Tacke, Chen & Manley, 1997). It has also been shown that hyper and hypo-phosphorylation of SR proteins may cause loss of splicing function (Kanopka et al., 1998; Prasad et al., 1999; Sanford & Bruzik, 1999). Post-translational modification has been shown
to be the means by which SR protein activities are strictly controlled during early development in *Ascaris lumbricoides*, sex determination in *Drosophila*, and during adenovirus infection (Du, Melnikova & Gardner, 1998; Kanopka et al., 1998; Sanford & Bruzik, 1999). Some SR proteins cannot be detected in the fully unphosphorylated state in cells or tissues, and when unphosphorylated recombinant SR proteins are used in an in vitro splicing assay, they become rapidly phosphorylated (Hanamura et al., 1998).

SR proteins assist in the early stages of spliceosome assembly, especially during the formation of the E complex (Jamison et al., 1995; Kohtz et al., 1994; Staknis & Reed, 1994; Wu & Maniatis, 1993) through RS domain protein-protein interactions. For example, SR proteins simultaneously interact with both snRNP U1-70K and U2AF35 through the RS domain (Wu & Maniatis, 1993). SR proteins also help in recruitment of the tri-snRNP U4/U6·U5 (Roscigno & Garcia-Blanco, 1995; Tarn & Steitz, 1995) perhaps through interaction between the RS domain of the SR protein and the RS domain of snRNP U4/U6·U5-27K, although the precise mechanism is unknown. SR proteins binding to ESEs can also enhance the second catalytic step of splicing (Chew et al., 1999), and SR proteins also function to regulate alternative splice-site selection (Caceres et al., 1994; Ge & Manley, 1990; Krainer et al., 1990a). SF2/ASF was recently shown to be a proto-oncogene that can fully transform immortal cells (Karni et al., 2007).
Figure 3. SR proteins. The diagrams show a list of some SR proteins and features common to all SR proteins. The red ovals represent the RNA-recognition motif (RRM) and the green rectangles represent the arginine-serine (RS) domain.
B. hnRNP proteins

Heterogeneous nuclear ribonucleoproteins (hnRNPs) are trans-acting factors that bind to ESSs or ISSs, usually to inhibit splicing during regulated splicing events. There are some instances in which hnRNP proteins promote splicing instead of inhibiting it (Caputi & Zahler, 2002; Mayeda, Helfman & Krainer, 1993; Paradis et al., 2007). The most common features of the hnRNP proteins is that they all have two or more RNA-binding domains and an auxiliary domain believed to be responsible for protein-protein, RNA-protein, and single-stranded DNA-protein interactions. Most of these hnRNPs can also form homophilic interactions and heterophilic interactions with other hnRNPs (Cartegni et al., 1996; Cobianchi et al., 1988; Nadler et al., 1991).

Heterogeneous nuclear ribonucleoproteins were first discovered by a nuclear sub-fractionation technique that removes about 99% of the associated chromatin as a part of the hnRNA complex, a major group of chromatin-associated RNA-binding proteins (Herman, Weymouth & Penman, 1978). hnRNP proteins are some of the most abundant nuclear proteins, and one of the most abundant of them has about 60 million molecules per HeLa cell nucleus (Dreyfuss et al., 1993; Hanamura et al., 1998). hnRNPs associate with the RNA during transcription, and some of them remain with it by the time the RNA is exported to the cytoplasm (Dreyfuss et al., 1993). These hnRNPs, which shuttle between the nucleus and the cytoplasm have nuclear export sequences (NESs) (Dreyfuss et al., 1993). Also, those hnRNPs that do not shuttle have nuclear retention sequences (NRS) within their auxiliary domain, and are restricted to the nucleus. This, perhaps, is to prevent unprocessed and partially processed RNA from being translated in the cytoplasm, which could be detrimental to the cell (Nakielny & Dreyfuss, 1996).

There are other roles that hnRNP proteins play in gene expression. For instance, hnRNP A1 binds single-stranded DNA telomere repeats and regulates telomere length through stimulation of telomerase activity (Ishikawa et al., 1993; LaBranche et al., 1998; McKay & Cooke, 1992; Zhang et al., 2006). hnRNP K has been shown to interact with the transcription machinery and to regulate transcription (Du et al., 1998; Michelotti et al., 1996). Also, mRNA 3’-end formation, maintenance and polyadenylation have been linked to several hnRNP proteins, including hnRNP I, hnRNP H, and hnRNP Nab4p (Bagga, Arhin & Wilusz, 1998; Castelo-Branco et al., 2004; Kessler et al., 1997; Moreira et al., 1998). There is increased expression of hnRNP A1 and SR proteins in tumors and tumor cell lines (Karni et al., 2007).
1.4 Antagonism between positive and negative regulators

Antagonism between positive regulators (the SR proteins) and negative regulators (the hnRNP proteins) usually determines whether the splicing pattern of a particular gene is constitutive or alternative, depending on which splicing regulator prevails. If the positive regulator prevails, the splicing pattern is constitutive, and if it is the negative regulator that prevails, then the splicing pattern is alternative. The positive regulators, the SR proteins, are trans-acting factors that usually bind to cis-acting elements, ESEs, and the negative regulators, the hnRNP proteins, usually bind to other cis-acting elements, ESSs. An example of this antagonism between positive and negative regulators can be seen in HIV-1 tat exon 3 between an SR protein, SC35, and hnRNP A1. This antagonism determines the exclusion or the inclusion of that particular exon. In this case, if hnRNP A1 prevail over SC35, exon 3 is skipped, but if another SR protein, SF2/ASF, replaces SC35; SF2/ASF will prevail over hnRNP A1 causing the inclusion of exon 3 (Zhu et al., 2001).

1.5 References


Zhu, J., Mayeda, A. & Krainer, A. R. (2001). Exon identity established through differential antagonism between exonic splicing silencer...
Chapter 2

Mechanism of hnRNP A1 Cooperative Binding
2.1 Abstract

hnRNP A1 binds to RNA in a cooperative manner. Initial hnRNP A1 binding to an exonic splicing silencer (ESS) at the 3’ end of HIV-1 tat exon 3, which is a high-affinity site, is followed by cooperative spreading in a 3’ to 5’ direction. As it propagates towards the 5’ end of the exon, hnRNP A1 antagonizes binding of a serine/arginine (SR) protein to an exonic splicing enhancer (ESE), thereby inhibiting splicing at that exon’s alternative 3’ splice site. Tat exon 3 and the preceding intron of HIV-1 pre-mRNA can fold into an elaborate RNA secondary structure, which could potentially influence hnRNP A1 binding. We report here that hnRNP A1 binding to RNA and splicing repression can occur on an unstructured RNA. Moreover, hnRNP A1 can effectively unwind an RNA hairpin upon binding. We further show that hnRNP A1 can also spread in a 5’ to 3’ direction, although when initial binding takes place in the middle of an RNA, spreading proceeds preferentially in a 3’ to 5’ direction. Finally, when two distant high-affinity sites are present, they can facilitate cooperative spreading of hnRNP A1 between the two sites.

2.2 Introduction

Splicing can be subdivided into constitutive and alternative. Constitutive splicing is the removal of introns by joining together all the adjacent exons in the order of their arrangement, without skipping any exon. In constitutive splicing, a single protein is produced from a single pre-mRNA, regardless of where and when the gene is expressed. In alternative splicing, variable use of splice sites allows two or more mature mRNAs to be generated from the same pre-mRNA. For example, an entire
exon or part of an exon can be included or skipped in different spliced mRNAs. Alternative splicing is a prevalent way by which many eukaryotes diversify the number of proteins produced from a single pre-mRNA transcript (Smith, Query & Konarska, 2008; Wang & Burge, 2008).

Heterogeneous nuclear ribonucleoproteins (hnRNPs) are trans-acting factors that bind to ESSs or ISSs, usually to inhibit splicing during regulated splicing events. There are some instances in which hnRNP proteins promote splicing instead of inhibiting it (Caputi & Zahler, 2002; Mayeda, Helfman & Krainer, 1993; Paradis et al., 2007). The most common feature of hnRNP proteins is the presence of two or more RNA-binding domains and an auxiliary domain believed to be responsible for protein-protein, RNA-protein, and single-stranded DNA-protein interactions. Most of these hnRNPs can also form homophilic interactions and heterophilic interactions with other hnRNPs (Cartegni et al., 1996; Cobianchi et al., 1988; Nadler et al., 1991). One of the most abundant hnRNPs is hnRNP A1 (Dreyfuss et al., 1993; Hanamura et al., 1998). hnRNP A1 has been implicated in many alternative splicing events in human and several other eukaryotes (Abdul-Manan & Williams, 1996; Amendt, Si & Stoltzfus, 1995; Blanchette & Chabot, 1999; Burd & Dreyfuss, 1994; Caputi et al., 1999; Hua et al., 2008; Mayeda & Krainer, 1992; Zhu, Mayeda & Krainer, 2001). Human hnRNP A1 is a 320-amino-acid protein, of which the 196-amino-acid N-terminal domain comprises two RNA-recognition motifs (RRMs) (Figure 4) (Maris, Dominguez & Allain, 2005). The 124-amino-acid C-terminal domain is glycine-rich (Figure 4) and is believed to be responsible for cooperative binding, leading to repression of splicing (Ding et al., 1999b; Shamoo et al., 1997). At present, there are no available structures of intact hnRNP A1, but there are crystal structures of its N-terminal domain spanning RRM1 and RRM2, which is known as UP1 (unwinding protein 1) (Figure 4) (Ding et al., 1999b; Shamoo et al., 1997; Vitali et al., 2002; Xu et al., 1997).
Figure 4 Structure of hnRNP A1 showing the domains. UP1 is a proteolytic cleavage of hnRNP A1 (as indicated).

The manner in which hnRNP A1 controls alternative splicing is still not fully understood. A study from our lab about splicing of exon 3 of the HIV-1 tat pre-mRNA showed an antagonistic effect of an exonic splicing silencer element, ESS3, mediated by hnRNP A1, vis-à-vis another cis-acting splicing regulatory element, known as an exonic splicing enhancer (ESE) (Zhu et al., 2001). ESEs enhance splicing or promote inclusion of a particular exon through the binding of one or more activator proteins, such as members of the serine/arginine (SR) family, which in turn recruit other components of the splicing machinery to the 5' and 3' splice sites (Huang & Steitz, 2005; Lin & Fu, 2007). SR proteins have one or two RRMs at their N-terminus, which interact with the RNA (Krainer, Conway & Kozak, 1990; Mayeda & Krainer, 1992; Tange & Kjems, 2001; Zhu & Krainer, 2000). The C-terminal domain of each SR protein comprises a highly conserved serine/arginine-rich (RS) domain (Krainer et al., 1990; Mayeda & Krainer, 1992; Tange & Kjems, 2001; Zhu & Krainer, 2000; Zhu et al., 2001); however, this domain is not always necessary for splicing (Shaw et al., 2007; Zhu & Krainer, 2000). SR proteins are important for the recognition of splice sites, and act at the earliest stages of spliceosome assembly, as well as at later stages of splicing (Krainer et al., 1990; Mayeda & Krainer, 1992; Tange & Kjems, 2001; Zhu et al., 2001). SR proteins have other functions in splicing and gene expression, besides binding to ESEs, and they are essential for constitutive splicing (Huang & Steitz, 2005). Even in the case of introns with strong splice sites, in which an ESE might not be required, SR proteins are essential for recognition of the splice sites and recruitment of the splicing machinery (Eperon et al., 2000; Huang & Steitz, 2005; Lin & Fu, 2007; Tange & Kjems, 2001; Zhu & Krainer, 2000).

The above-mentioned study showed that initial high-affinity binding of hnRNP A1 to ESS3 is followed by its cooperative spreading along tat
exon 3, which allows hnRNP A1 to displace the SR protein SC35 from its cognate ESE, thereby preventing splicing of tat exon 3 (Zhu et al., 2001). This same study also showed that when another SR protein, SF2/ASF, binds to its cognate ESE, hnRNP A1 cannot effectively displace it, and therefore, there is inclusion of tat exon 3 (Zhu et al., 2001). The net effect depends in part on the strength of the SR protein interaction with its cognate ESE, and presumably on the nuclear abundance of particular SR proteins and hnRNP A1 in a given cell type.

There is increased expression of hnRNP A1 or SR proteins in some tumors and tumor cell lines, as compared to normal cells and tissues (Ghigna et al., 1998; Karni et al., 2007; Perrotti & Neviani, 2007). Putting all this information together presents a strong case for studying how cooperative binding of hnRNP A1 leads to alternative splicing of a specific exon. Understanding cooperative binding of hnRNP A1 in the context of HIV-tat and other model substrates is expected to shed light on the mechanisms of alternative splicing in general.

The present study addresses the mechanism of hnRNP A1 cooperative binding. We show that hnRNP A1 cooperative binding results in unwinding of RNA secondary structure. After binding to a high-affinity site, hnRNP A1 spreads preferentially, though not exclusively, in a 3’ to 5’ direction, and can displace other bound proteins from the RNA to repress splicing.

2.3 Results

2.3.1 Cooperative binding of hnRNP A1 does not require RNA secondary structure.
Inhibition of splicing of exon 3 of an HIV-1 Tat23 mini-gene occurs through cooperative binding of hnRNP A1, such that multiple molecules bind by spreading from a high-affinity binding site (ESS3) at the 3’ end of the RNA towards the 5’ end (Zhu et al., 2001). The tat pre-mRNA can adopt an intricate secondary structure in solution (Damgaard, Tange & Kjems, 2002; Marchand et al., 2002), and it has been proposed that hnRNP A1 binding and silencing involves cooperative binding to these structured regions, rather than spreading along single-stranded RNA (Damgaard et al., 2002; Marchand et al., 2002). However, UP1 (unwinding protein 1), as its name indicates, can unwind RNA or DNA secondary structures (Herrick & Alberts, 1976; Zhang et al., 2006); hnRNP A1 facilitates annealing of complementary nucleic acid strands below their Tm (melting temperature); on the other hand, when hnRNP A1 binds to duplex DNA, it lowers the Tm of the duplex, thereby facilitating its unwinding; and at a temperature above the new Tm, hnRNP A1 can also maintain an equilibrium between single- and double-stranded DNA (Pontius & Berg, 1990; Pontius & Berg, 1992). However, hnRNP A1 had not been shown to be capable of unwinding RNA secondary structure.

To address the potential involvement of RNA secondary structure in hnRNP A1 cooperative binding, we generated by in vitro transcription RNA comprised mainly of oligo U tracts, with $^{32}$P-labeled C at every fifth nucleotide position. We chose this nucleotide composition because, hnRNP A1 has low affinity for poly U and poly C (Abdul-Manan & Williams, 1996). Near the 3’ end of the RNA, we placed a high-affinity hnRNP A1 binding site, UAGGGU, as determined by SELEX (Burd & Dreyfuss, 1994) (Figure 5A). Based on its composition and sequence, this RNA cannot form secondary structures, at least by conventional base pairing. To reduce other potential higher-order structures, the RNA was denatured at 95 °C and rapidly cooled before incubation with recombinant hnRNP A1 at different concentrations. The complex formed between this RNA and hnRNP A1 was subjected to UV cross-linking, followed by digestion with ribonucleases A and T1, separation by SDS-PAGE, and detection by autoradiography (Figure 5B). The transfer of label to hnRNP A1 can be detected after nuclease digestion, because the protein spreads along the RNA from the high-affinity site (Zhu et al., 2001). This was confirmed by the reduction in labeled hnRNP A1 when the high-affinity site was mutated at a single nucleotide, from UAGGGU (WT) to UUGGGU (MUT). We conclude that unstructured RNA is compatible with cooperative binding of hnRNP A1.
To verify that the SELEX winner UAGGGU can act as an exonic splicing silencer (ESS), we constructed a β-globin minigene, Nβ2, comprising the last 101 nucleotides of exon 1 and the first 101 nucleotides of exon 2, and inserted UAGGGU at the 3’ end of exon 2, followed by a BamH1 site (Figure 5C). A similar control minigene, Nβ3, has a single point mutation changing UAGGGU to UCUGGU, which abrogates hnRNP A1 binding (as does UUGGGU; see Figure 5B and Discussion). Finally, Nβ1 is the parental minigene without an inserted hexamer. Labeled pre-mRNAs transcribed from these minigenes were spliced in HeLa cell cytoplasmic extract (S100) complemented with recombinant SC35, in the presence or absence of recombinant hnRNP A1 (which is limiting in S100 extract; ref. 57) (Figure 5D). The results show that splicing repression requires an intact ESS (cf. lanes 1-3, right panel) and a sufficient amount of hnRNP A1 (cf. lanes 2 in right and left panels). Moreover, the A to C point mutation at position 2 of the ESS is sufficient to abolish splicing silencing (cf. lanes 2 and 3, right panel), and another mutation in the ESS, position 2 A to U, also abolishes hnRNP A1 cooperative binding (Figure 5B).

**Figure 5.** hnRNP A1 cooperative binding does not require RNA secondary structure. (A) Sequences of wild-type (WT) and mutant (MUT) RNAs for UV-crosslinking experiments. The underlined hexanucleotide is a high-affinity hnRNP A1 binding site: UAGGGU is the wild-type version, and UCUGGU is the inactive, mutant version, with the mutated nucleotide shown in italics. The radiolabeled cytidines incorporated by in vitro transcription are indicated in bold italics. (B) UV crosslinking with WT and MUT RNAs from (A) in the presence of increasing concentration of recombinant hnRNP A1. The crosslinked products were digested with RNases A and T1, separated by SDS-PAGE, and detected by autoradiography. Band intensities were measured on a

---

**Figure 5 continued:**

- **A**
  - WT: 5’GGGAGUUUUCUUCUCUCUCUCUCUCUCUCUCUCUGGGAUGGGU3’
  - MUT: 5’GGGAGUUUUCUUCUCUCUCUCUCUCUCUCUCUCUGGGAUCUGGU3’

- **B**
  - WT: UAGGGU
  - MUT: UCUGGU
  - 0.5 1 2 4
  - 0.5 1 2 4
  - Relative Intensity

- **C**
  - Nβ1
  - Nβ2
  - Nβ3

- **D**
  - S100+SC35
  - +hnRNP A1
  - M 1 2 3 % 1 2 3 Splicing
phosphorimager, and normalized values relative to the lowest band intensity are shown below the gel. (C) β-globin minigene transcripts for in vitro splicing assays. The pre-mRNAs comprise 108-nt of exon 1, the 130-nt first intron, and 108-nt of exon 2. Nβ2 has an additional 6-nt ESS at the 3’ end; Nβ3 has a mutant version of the ESS (ESSm). (D) Splicing of the pre-mRNAs from (C) in HeLa S100 extract complemented with SC35, in the presence or absence of 7.5 pmol of hnRNP A1. Splicing efficiency (mRNA/(pre-mRNA+mRNA)x100%) is shown below the autoradiogram. Quantitation of the data is based on each representative experiment shown in the figure; consistent trends were observed in repeat experiments (3 times for panel B and 10 times for panel D).

2.3.2 hnRNP A1 unwinds RNA secondary structure.

To further test whether or not cooperative binding of hnRNP A1 involves RNA secondary structure, we took advantage of an RNA with known secondary structure, namely a natural hairpin that binds bacteriophage MS2 coat protein (Graveley & Maniatis, 1998; LeCuyer, Behlen & Uhlenbeck, 1995). We inserted the MS2 hairpin in the middle of exon 2 of the β-globin minigene and included the ESS at the 3’ end. As a control, we inserted a hairpin with deletion of a single bulged nucleotide, to abolish MS2 protein binding (Graveley & Maniatis, 1998) (constructs MS2 and MS2* in Figure 6A). We expected that tight binding by MS2 to the wild-type construct, but not to the mutant construct, would block hnRNP A1 propagation along the exon, and therefore prevent splicing repression. In addition, omitting the MS2 coat protein should allow us to determine whether both hairpins would block the spreading of hnRNP A1.

The results we obtained were unexpected: we observed inhibition of splicing in the construct with the MS2 hairpin loop and the ESS, in the presence of MS2 coat protein (Figure 6B, cf. lanes 14 and 15 with lanes 17 and 18). There are several possible explanations for this result: first, RNA secondary structure may actually facilitate cooperative binding of hnRNP A1 (Damgaard et al., 2002; Marchand et al., 2002) despite the presence of bound MS2 coat protein; second, hnRNP A1 may unwind the
hairpin, displacing the tightly bound MS2 coat protein and cooperatively spread along the exon to repress splicing; third, bound MS2 coat protein permits or perhaps facilitates cooperative binding of hnRNP A1, although this seems improbable; and fourth, in spreading along the exon, hnRNP A1 may somehow bypass the hairpin with bound MS2 coat protein. The splicing assay in this section was done at least seven times with very similar results, and the gel shown represents the best of the seven trials without any statistical values.

To distinguish among these possibilities, we used GST pulldowns to measure whether MS2 coat protein is displaced by hnRNP A1 cooperative binding. We made four artificial RNA transcripts composed mainly of oligo-U with 32P-labeled C every fifth nucleotide, with an MS2 hairpin in the middle and a high-affinity hnRNP A1 binding site at the 3’ end (Figure 6C). Each construct is either 90 or 91 nucleotides long, depending on whether it has an MS2 or MS2* version of the hairpin. Each RNA construct was denatured at 95 °C and allowed to refold at room temperature before adding GST-MS2 protein. The results shown in Figure 6D clearly demonstrate that hnRNP A1 displaces bound GST-MS2 protein, presumably by unwinding the stem-loop and/or by physical displacement (Figure 6D, left panel).

As the amount of hnRNP A1 protein increases, the amount of GST-MS2 protein displaced increases, and this effect largely depends on the initial binding of hnRNP A1 to the high-affinity site (cf. the first five lanes (PUMS2ESS) WT with the last five lanes (PUMS2ESSm) in which a single point mutation abrogates the high-affinity hnRNP A1 binding site). The control RNA with the MS2* hairpin failed to bind MS2 coat protein, as expected (data not shown). A similar experiment was done with the same four RNA constructs and UP1 protein, which cannot undergo cooperative binding (Zhu et al., 2001). As expected, UP1 was unable to displace GST-MS2 protein (Figure 6D, right panel).
2.3.3 Cooperative binding of hnRNP A1 can also proceed from 5’ to 3’ to inhibit splicing.
Previous studies of hnRNP A1 cooperative binding focused on oligomerization in a 3’ to 5’ direction, after initial binding to an ESS at the 3’ end of a pre-mRNA (Damgaard et al., 2002; Del Gatto-Konczak et al., 1999; Marchand et al., 2002; Zhu et al., 2001). To determine if hnRNP A1 can also spread in a 5’ to 3’ direction, we generated an artificial RNA comprising mainly oligo U tracts, with labeled $^{32}$P C at every fifth position, and a high-affinity hnRNP A1 SELEX winner sequence UAGGGU (Burd & Dreyfuss, 1994), at the 5’ end (Figure 7A). We also made a control RNA with a mutated hnRNP A1 binding sequence, UUGGGU. UV cross-linking (Zhu et al., 2001) of these RNAs after incubation with increasing concentration of recombinant hnRNP A1 was followed by RNase digestion, SDS-PAGE, and autoradiography (Figure 7B). This experiment shows that cooperative binding of hnRNP A1 can proceed in a 5’ to 3’ direction.

To test the effect of 5’ to 3’ hnRNP A1 cooperative binding on splicing, we designed short β-globin-derived minigene constructs with a 64-nt exon 1 and a 109-nt exon 2. We engineered a high-affinity hnRNP A1 binding sequence, UAGGGU (ESS), at the 5’ of exon 1 by PCR; controls including a mutant hnRNP A1 motif, UCGGGU (ESSm), and a construct without the hnRNP A1 binding site were similarly made by PCR (Figure 7C). In vitro splicing of pre-mRNA transcribed from these constructs in S100 extract complemented with SC35, with or without addition of hnRNP A1, is shown in Figure 7D. Cooperative binding of hnRNP A1 propagating in a 5’ to 3’ direction in exon 1 inhibited splicing (cf. lanes 4 to 5, right panel and lanes 4 and 5, left panel).
Figure 7. hnRNP A1 cooperative binding spreading from the 5' end to the 3' end of exon 1 inhibits splicing. (A) Sequences of synthetic wild-type (WT) and mutant (MUT) RNAs for UV-crosslinking experiments. The underlined hexanucleotide is a high-affinity hnRNP A1 binding site: UAGGGU is the wild-type version, and UUGGGG is the inactive, mutant version, with the mutated nucleotide shown in italics. The radiolabeled cytidines are indicated in bold italics. (B) UV crosslinking with WT and MUT RNAs from (A) in the presence of increasing concentration of recombinant hnRNP A1. Detection and quantitation of the crosslinked products was as in Figure 1B. (C) NSβ-globin minigene transcripts for in vitro splicing assays. The pre-mRNAs comprise 58-nt of exon 1, the 130-nt first intron, and 108-nt exon 2. NSβ2, 3, 4, and 5 have in addition a 6-nt ESS or mutant ESSm at either the 5' end of exon 1 or the 3' end of exon 2; NSβ6 has the 6-nt ESS at both the 5' end of exon 1 and the 3' end of exon 2. (D) In vitro splicing of pre-mRNAs from (C) in S100 extract complemented with SC35, in the presence or absence of 7.5 pmol of hnRNP A1. Splicing efficiency (calculated as in Figure 5D) is shown below the autoradiogram. Quantitation of the data is based on each representative experiment shown in the figure; consistent trends were observed in repeat experiments (3 times for panel B and 4 times for panel D).

2.3.4 hnRNP A1 preferentially spreads in a 3’ to 5’ direction.
We next sought to determine whether hnRNP A1 can undergo cooperative binding with bidirectional spreading. To this end, we generated two RNAs with an hnRNP A1 high-affinity binding site in the middle (Figure 8A). In the first construct, four nucleotides (cytosines) at every fifth position 3' of the hnRNP A1 binding site were radiolabeled, whereas the sequences 5' of the binding site were unlabeled. In the second construct, the labeled and unlabeled regions were reversed. Control substrates with a mutant hnRNP A1 binding site were also generated. UV crosslinking, RNase digestion, and SDS-PAGE analysis were carried out as in Figure 1. The ratio of WT over MUT intensities was greater for the 3' to 5' substrate (Figure 8C) than for the 5' to 3' substrate (Figure 8B), indicating preferential spreading of hnRNP A1 towards the 5' end.

Figure 8. Directionality of hnRNP A1 cooperative spreading. (A) Sequences of synthetic RNA transcripts for UV-crosslinking experiments. The radiolabeled cytidines are indicated in bold italics. The underlined hexanucleotide is a high-affinity hnRNP A1 binding site (ESS). The top transcript (5' to 3' Binding), has radiolabeled cytidines to the right of the ESS, whereas the bottom transcript (3' to 5' Binding), has them to the left of the ESS. (B) UV crosslinking with WT transcript (5' to 3' Binding) and its ESSm (MUT) version in the presence of increasing recombinant hnRNP A1. Detection and quantitation of the crosslinked products was as in Figure 5B. (C) As in panel B but with 3' to 5' Binding WT and MUT transcripts. Quantitation of the data is based on each representative experiment shown in the figure; consistent trends were observed in repeat experiments (3 times for panel B and 3 times for panel C).
2.3.5 Determining the extent of spreading of hnRNP A1 along the RNA.

The pre-mRNA we used to test for hnRNP A1 oligomerization by UV crosslinking has six labeled C nucleotides (Figure 5A). To determine more precisely how far hnRNP A1 spreads from the site of initial binding, we generated transcripts for crosslinking with fewer labeled Cs placed at different positions, to see if we could still detect label transfer, reflecting cooperative binding (Figure 9A). When the first two labeled Cs upstream of the high-affinity binding site were substituted with unlabeled Gs, we still detected label transfer to hnRNP A1 (Figure 9B, left panel) indicating that cooperative binding extends beyond ~20 nucleotides. Similarly, when the next two labeled C nucleotides were also substituted by unlabeled Gs, we continued to detect a signal (Figure 9B, right panel), indicating cooperative binding beyond ~30 nucleotides. Finally, we prepared a substrate by $^{32}$P 5'-end-labeling an otherwise unlabeled RNA transcript (Figure 9C), and again, we detected cooperative binding by comparing the wild type (WT) with the mutant (MUT1 or MUT2) transcripts (Figure 9D, top and bottom panels). As expected, the signals became progressively weaker as transcripts with fewer labeled nucleotides were analyzed. We conclude that multiple molecules of hnRNP A1 bind consecutively along the RNA, all the way to its 5' end.

To address the cooperative spreading of hnRNP A1 using a different technique, we carried out hydroxyl-radical footprinting using the first two $^{32}$P 5'-end-labeled RNAs in Figure 9C. Figure 9E, left panel, shows the footprinting results. With increasing recombinant hnRNP A1, the region protected by hnRNP A1 increased (cf. WT on lanes 2 to 5 with MUT on lanes 6 to 9). Figure 9E, right panel, shows RNase A footprinting, which gives consistent results (cf. WT on lanes 2 to 5 with MUT on lanes 7 to 10). Both footprinting methods show that the entire length of the 5'
labeled WT RNA is protected by cooperative binding of hnRNP A1. These results indicate that cooperative binding of hnRNP A1 to RNA resembles ‘beads on a string’, and does not require RNA secondary structure. As shown above, such structures, when present, can actually be unwound by hnRNP A1.

Figure 9. Mapping the extent of hnRNP A1 cooperative spreading along the RNA. (A) Sequences of synthetic RNA transcripts for UV-crosslinking experiments. The radiolabeled cytidines are indicated in bold italics. (B) UV crosslinking of RNA transcripts in (A) and the corresponding ESSm controls, in the presence of increasing recombinant hnRNP A1. Left panel: crosslinked products after RNAase digestion, of the top transcript in (A) and its ESSm counterpart. Right panel: Idem for the bottom transcript in (A) and its ESSm counterpart. Band intensities were measured on a phosphorimager, and normalized values are shown below the gel. (C) Sequences of 5’-end labeled synthetic RNA transcripts for UV-crosslinking and footprinting experiments. The underlined hexanucleotide is a high-affinity hnRNP A1 binding site: UAGGGU is the wild-type version (WT),
UUGGGU is the inactive, mutant version 1 (MUT1), and UUUGGU is the inactive, mutant version 2 (MUT2), with the mutated nucleotide shown in italics. (D) UV crosslinking with WT and MUT1 and MUT2 RNAs from (C) in the presence of increasing concentration of recombinant hnRNP A1. Detection and quantitation of the crosslinked products was as in Figure 5B. (E) Footprinting assays with the first two RNA transcripts in C (WT and MUT1); the left panel shows a hydroxyl-radical footprinting assay in the presence of increasing recombinant hnRNP A1, and the right panel is an RNase A footprinting assay. (M: molecular weight markers). Quantitation of the data is based on each representative experiment shown in the figure; consistent trends were observed in repeat experiments (3 times for panels B and D, and 2 times for panel E).

2.3.6 “Cross-talk” between hnRNP A1 molecules bound at distant sites.

Finally, we investigated whether distant high-affinity hnRNP A1 binding sites can influence how hnRNP A1 binds to each site and subsequently spreads. Relevant to this, hnRNP A1 was reported to dimerize upon binding to distant sites, apparently looping out the RNA between the two sites (Nasim et al., 2002). We generated five RNA constructs (Figure 10). First, we placed two identical ESSs at different positions along the RNA constructs (Figure 10A): the first construct, XT1, has the two ESSs juxtaposed, separated by only two nucleotides, and placed at the 3′ end of the RNA; the second construct, XT2, has one ESS at the 3′ end and the other in the middle of the RNA; the third construct, XT3, has one ESS at the 3′ end and the other at the 5′ end of the RNA; the fourth construct, XT3m1, has a mutant ESS (ESSm) at the 5′ end of the RNA (Figure 10C); and the fifth construct, XT3m2, has a mutant ESS (ESSm) at the 3′ end (Figure 10C). UV cross-linking of each of these constructs in the presence of increasing amounts of recombinant hnRNP A1 was compared with that of an RNA with a single ESS at the 3′ end (WT) (see Figure 5A).

When the two binding sites were separated by only two nucleotides, there was no apparent cross-talk between the two sites, i.e., no additive or synergistic effect compared to the control WT RNA (Figure 10B, left panel, cf. WT with XT1). When the distance separating the two high-affinity
binding sites was greater, the signal relative to the WT RNA increased (Figure 10B, middle panel, cf. WT with XT2), indicating synergy, or cross-talk, between the sites. With the ESS at both ends of the RNA, the signal increased even further (Figure 10B, right panel, cf. WT with XT3). When the ESS at either end of the RNA construct was inactivated by a point mutation (Figure 10C), cooperative-binding-dependent cross-talk was lost (Figure 6D, cf. XT3 with XT3m1, and also XT3 with XT3m2).

We note that the type of cross-talk shown in this experiment is cooperative-binding-dependent, and differs from the mechanism proposed by Nasim et al (Nasim et al., 2002): if the RNA between the two distant high-affinity sites is looped out, label transfer would not occur, because all the radiolabeled nucleotides were present along the RNA sequences between the two ESSs at the ends.

Figure 10. “Cross-talk” between hnRNP A1 molecules bound at distant sites. (A) Sequences of synthetic RNA transcripts for UV-crosslinking experiments. The radiolabeled cytidines are indicated in bold italics. The underlined hexanucleotide is a high-affinity hnRNP A1 binding site (ESS). (B)
UV crosslinking with RNA transcripts from (A) and WT control from Figure 1A, in the presence of increasing recombinant hnRNP A1. Detection and quantitation of the crosslinked products was as in Figure 1B. The position of the ESS in each of the RNAs is indicated by a dark line. (C) Sequences of synthetic RNA transcripts for UV-crosslinking experiments. The radiolabeled cytidines are indicated in bold italics. The underlined hexanucleotide is a high-affinity hnRNP A1 binding site: UAGGGU is the wild-type version, and UUGGGU is the inactive, mutant version, with the mutated nucleotide shown in italics. (D) UV crosslinking with RNA transcripts from (C) and the XT3 control from (A), in the presence of increasing recombinant hnRNP A1. Detection and quantitation of the crosslinked products was as in Figure 5B. The position of the ESS in each of the RNAs is indicated by a dark line; a mutant ESS is indicated by an *. Quantitation of the data is based on each representative experiment shown in the figure; consistent trends were observed in repeat experiments (2 times for panel B and 2 times for panel D).

2.4 Discussion

We have demonstrated that RNA secondary structure is not required for hnRNP A1 cooperative binding to RNA, in contrast to suggestions from previous studies of hnRNP A1 binding to the HIV-1 tat pre-mRNA, which is highly structured in solution (Damgaard et al., 2002; Marchand et al., 2002). We found that hnRNP A1 can unwind RNA secondary structure in a cooperative-binding-dependent manner (Figure 6D). This result is consistent with hnRNP A1’s established properties as a single-stranded RNA/DNA binding protein that can coat the entire length of a polynucleotide (Cartegni et al., 1996; Cobianchi et al., 1988; Ding et al., 1999a). In vivo, this type of binding could play a multitude of roles in co-transcriptional and post-transcriptional RNA processing, including splice-site recognition, alternative splicing regulation, mRNA susceptibility to ribonucleases, nuclear export of mature mRNA, etc., as well as in telomere-length regulation (LaBranche et al., 1998; Zhang et al., 2006).

We showed that displacement of GST-MS2 protein bound to a hairpin and unwinding of this hairpin structure by hnRNP A1 require cooperative binding. Thus, UP1 had little or no activity in the GST-MS2-displacement and hairpin-unwinding assays (Figure 6D). This is consistent
with UP1 lacking the C-terminal glycine-rich domain, which is necessary for cooperative binding and splicing silencing (Ding et al., 1999b; Mayeda et al., 1994; Zhu et al., 2001). In addition, when we prevented the initial binding of hnRNP A1 by a point mutation in the high-affinity binding site, the protein could no longer displace bound GST-MS2 protein or unwind a hairpin.

We further showed that a 6-nt hnRNP A1 SELEX winner sequence (UAGGGU) has ESS activity, and that a single point mutation in this sequence is enough to disrupt hnRNP A1 cooperative binding and splicing silencing (Figure 5B and D). Note that we used two different hnRNP A1 binding site mutants, UUGGGU and UCGGGU; the latter disrupted hnRNP A1 binding to a greater extent than the former (data not shown). However, we used UUGGGU for binding and crosslinking experiments to avoid introducing a labeled C nucleotide into the hnRNP A1 binding site. On the other hand, because this was not a consideration for the splicing experiments, we used the more disruptive UCGGGU mutation for the splicing assays. Additional mutations we tested that also effectively disrupted the hnRNP A1 binding site were UACGGU and UAUGGU (data not shown).

Cooperative binding by hnRNP A1 was shown by our lab to spread from the 3’ end of an HIV-1 RNA towards the 5’ end of the exon, and to inhibit splicing by blocking an SC35-dependent ESE (Zhu et al., 2001). However, it was not known whether cooperative binding of hnRNP A1 can also proceed in a 5’ to 3’ direction and likewise inhibit splicing. Here, we observed that 5’ to 3’ cooperative spreading does occur, but appears to be considerably weaker than 3’ to 5’ spreading (Figure 7). We generated β-globin minigene derivatives with two exons of the same length (101 nt), and with the identical 6-nt ESS at the 5’ end of exon 1 in one construct, and at the 3’ end of exon 2 in the other construct. Using these pre-mRNAs, we observed strong inhibition of splicing in vitro for the pre-mRNA with the ESS at the 3’ end of exon 2 (Figure 8D), whereas splicing of the pre-mRNA with the ESS at the 5’ end of exon 1 was unaffected (data not shown). However, when we reduced the size of exon 1 with the ESS at the 5’ end to 64 nt, splicing was strongly inhibited (Figure 7D). This inhibition of splicing can be attributed to hnRNP A1 cooperative binding, as strong splicing inhibition depended on addition of recombinant hnRNP A1 (Figure 7D).
We also investigated whether cooperative binding of hnRNP A1 could proceed simultaneously in both directions. To this end, we placed the high-affinity binding site in the middle of an exon, and found that hnRNP A1 spreading proceeded preferentially from 3’ to 5’ (Figure 8B and C). A two-fold reduction in the concentration of RNA and protein was enough to abrogate 5’ to 3’ cooperative binding (data not shown). In light of this evidence, we conclude that hnRNP A1 5’ to 3’ cooperative binding is weaker than 3’ to 5’ binding.

Figure 11 shows our model for hnRNP A1 cooperative binding. hnRNP A1 can displace a protein bound to a secondary structure that interrupts the path of hnRNP A1 spreading. Moreover, hnRNP A1 unwinds the structure to then spread further and displace bound SC35 from an ESE. In a similar experiment, hnRNP A1 cooperative binding was unable to displace SF2/ASF from its ESE to inhibit splicing (data not shown); this is consistent with the tighter binding of SF2/ASF to its cognate ESE, compared to SC35 (Zhu et al., 2001).

A form of cross-talk or communication between two hnRNP A1 molecules bound at distant sites has been described (Nasim et al., 2002). This cross-talk allows the skipping of an exon between the two flanking intronic binding sites, through protein-protein interaction between hnRNP A1 molecules bound at these sites causing the exon to loop out. A similar looping out may also occur within a long intron, thereby increasing the efficiency of splicing between two distant splice sites (Nasim et al., 2002). Here we also investigated if there is cross-talk between two molecules of hnRNP A1 bound at distant high-affinity sites. Our results are consistent with a kind of cross-talk that does not involve looping out of the RNA. We found that when two high-affinity hnRNP A1 binding sites are juxtaposed, the extent of hnRNP A1 cooperative spreading towards the 5’ end of the RNA is similar to that observed with a single site (Figure 10). In contrast, as the distance between the two sites increases, the extent of cooperative binding increases and is maximal when the two high-affinity sites are placed at both ends of the RNA (Figure 10). We did not observe looping out of the RNA between the two high-affinity sites; in the context of our experiments, looping out would not have resulted in label transfer to hnRNP A1, as all the labeled nucleotides were placed between the two high-affinity sites.

We termed the kind of interaction between two hnRNP A1 sites observed here cooperative-binding-dependent cross-talk. With the two
hnRNP A1 binding sites placed at the 5' and 3' ends of the RNA respectively, hnRNP A1 binding initially at the 5' site would spread towards the 3' end, and simultaneously, hnRNP A1 binding initially at the 3' end would spread towards the 5' end. Convergent spreading would increase the rate at which the gap between the two binding sites is filled with hnRNP A1 molecules, compared to a single initial binding site (Figure 10b, right panel). The looping model (Nasim et al., 2002) and the cross-talk model reported here may each apply in different situations, although what pre-mRNA contexts or cellular conditions determine one or the other mode of binding remains unknown.

The results presented here indicate that hnRNP A1 can unwind an RNA hairpin, even when the hairpin is protected by a tightly bound protein. However, it is possible that more extensive secondary/tertiary structures and/or very tightly bound proteins could be more effective at blocking hnRNP A1 propagation, compared with the MS2 hairpin with or without bound MS2 coat protein.

In short, we have described the features of hnRNP A1 cooperative binding. This cooperative binding, as shown in Figure 11, unwinds RNA secondary structure, and preferentially spreads in a 3' to 5' direction to displace SR proteins bound at an ESE, thereby inhibiting splicing. 5' to 3' cooperative spreading of hnRNP A1 appears to be less robust, but within certain distance constraints, it may also be sufficient to unwind RNA secondary structure, displace bound SR proteins, and/or displace U1 snRNP from a 5' splice site to inhibit splicing.
2.5 Methods and Materials

2.5.1 Transcripts

All pre-mRNA transcripts for in vitro splicing were 5’ capped and labeled by in vitro transcription in the presence of [α-32P]-UTP from PCR templates with a T7 phage promoter (Mayeda & Krainer, 1999a). The PCR primers to generate transcription templates for all constructs used for in vitro splicing are listed in Table 1. The template for PCR was the
linearized plasmid pSP64-HβΔ6 (Krainer et al., 1984) except for constructs NβMS2ESS, NβMS2ESSm, NβMS2*ESS, and NβMS2*ESSm, whose PCR templates were NβMS2ΔESS and NβMS2*ΔESS respectively. All the model RNA transcripts were transcribed in the presence of [α-32P]-CTP from the corresponding antisense oligonucleotides with a T7 phage promoter annealed with a T7 sense oligonucleotide as described (Milligan & Uhlenbeck, 1989). All unlabeled model RNA transcripts were similarly transcribed from synthetic oligonucleotide templates, using a T7-MEGAscript kit (Ambion catalog #1354), followed by 5' end labeling with [γ-32P]-ATP (Romaniuk & Uhlenbeck, 1983). For antisense oligonucleotides corresponding to Figure 2C, see Table 2.

Table 1. List of DNA primers used to generate transcription templates by PCR.

<table>
<thead>
<tr>
<th>Names</th>
<th>Forward Primers</th>
<th>Reverse Primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB1</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>TAGGCAGACACCATGGTGCACC</td>
</tr>
<tr>
<td>NB2</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>CCGGGAATCCGTACCCTAATGAGCC</td>
</tr>
<tr>
<td>NB3</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>CTTACCCCTAATGAGGC</td>
</tr>
<tr>
<td>NB4</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>CTTACCCCTAATGAGGC</td>
</tr>
<tr>
<td>NB5</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>CTTACCCCTAATGAGGC</td>
</tr>
<tr>
<td>NB6</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>CTTACCCCTAATGAGGC</td>
</tr>
<tr>
<td>NSβ1</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>TAGGCAGACACCATGGTGCACC</td>
</tr>
<tr>
<td>NSβ2</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>CCGGGAATCCGTACCCTAATGAGGC</td>
</tr>
<tr>
<td>NSβ3</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>CCGGGAATCCGTACCCTAATGAGGC</td>
</tr>
<tr>
<td>NSβ4</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>CCGGGAATCCGTACCCTAATGAGGC</td>
</tr>
<tr>
<td>NSβ5</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>CCGGGAATCCGTACCCTAATGAGGC</td>
</tr>
<tr>
<td>NSβ6</td>
<td>GAATACAACCTGTGATACACTACATTA</td>
<td>CCGGGAATCCGTACCCTAATGAGGC</td>
</tr>
</tbody>
</table>

Table 2. List of antisense oligonucleotides used as transcription templates to generate the RNA transcripts showed in Figure 2C. (Underlined sequence: MS2 or MS* hairpin)

<table>
<thead>
<tr>
<th>Names</th>
<th>Anti-sense Oligos</th>
</tr>
</thead>
<tbody>
<tr>
<td>PUMS2ESS</td>
<td>GATCCGTACCCTAAAAAGAAAAGAAAAGAAAAAA</td>
</tr>
<tr>
<td>PUMS2ESSm</td>
<td>GATCCGTACCCGAAAAAGAAAAGAAAAGAAAAAAA</td>
</tr>
<tr>
<td>PUMS2*ESS</td>
<td>GATCCGTACCCTAAAAAGAAAAGAAAAGAAAAAA</td>
</tr>
</tbody>
</table>
2.5.2 Recombinant Proteins

Untagged human hnRNP A1 was expressed in E. coli and purified as described (Mayeda et al., 1994). Purified GST-MS2 protein expressed in E. coli was a gift from Zuo Zhang. Purified human SC35 expressed in baculovirus was a gift from Michelle Hastings. Purified UP1 expressed in E. coli was a gift from Qingshuo Zhang.

2.5.3 In Vitro Splicing Assays

S100 extract from HeLa cells was prepared as described (Mayeda & Krainer, 1999b). In vitro splicing reactions were carried out in a final volume of 12.5 µl with 15 fmol (1.15 nM) of $^{32}$P-labeled, $^7$CH$_3$-GpppG-capped T7 RNA transcripts, 35% (v/v) S100 extract, with a final concentration of 0.4 µM SC35, in the presence or absence of hnRNP A1 at a final concentration of 0.6 µM, and in the presence or absence of GST-MS2. All the $^{32}$P-labeled RNAs in Figure 6A were first incubated with GST-MS2 at a final concentration of 1.73 µM in standard splicing buffer (Mayeda & Krainer, 1999a) for 15 min at room temperature or at 30 °C before the addition of extract mix with or without hnRNP A1, and further incubation at 30 °C for 2 hr as described (Mayeda & Krainer, 1999a).
2.5.4 UV Crosslinking

All UV crosslinking assays were performed in a Spectronics XL1000 instrument at 0.48 J/cm² under splicing-reaction conditions as described (Zhu & Krainer, 2000; Zhu et al., 2001) except that the buffer was slightly modified by addition of final concentrations of 1 mg/ml heparin, 0.16 mg/ml yeast tRNA, and 0.11 mg/ml BSA with 8-32 nM ³²P-labeled RNA. Before the UV crosslinking, binding of hnRNP A1 to RNA was done as follows: (1) 0.5 µL of 40 mM MgCl₂, 0.1 µL of 20 mg/ml tRNA, 0.25 µL of 50 mg/mL heparin, 1.0 µL of 100-400 nM ³²P-labeled RNA and 1.9 µL of RNase-free H₂O were first incubated together at 95 °C for 5 minutes to unwind the RNA, and then placed on ice immediately, followed by addition of 0.132 µL of 10 mg/ml BSA, 0.625 µL of 40 mM Hepes-KOH pH 7.3 and 0.5 µL of 12.5 mM ATP/0.5 M creatine phosphate mix; (2) The extract mix contained variable concentration of hnRNP A1 in Buffer D with 100 mM KCl. Finally, 5 µL of Buffer mix was incubated together with 7.5 µl of Extract mix at 30 °C for 20 minutes. The concentration of each component in the final reaction volume was: 1.6 mM MgCl₂, 0.16 mg/mL tRNA, 1 mg/mL heparin, 8-32 nM of ³²P-labeled RNA, 0.11 mg/mL BSA, 2 mM Hepes-KOH pH 7.3, 0.5 mM ATP/20 mM creatine phosphate, and 60 mM KCl.

2.5.5 GST-MS2 Pulldowns
32P-labeled RNAs used in the GST-MS2 pulldowns were first incubated at 95 °C for 3-5 min and then allowed to refold at room temperature for 1-2 min in crosslinking buffer as described above. Renatured RNAs were then allowed to form complexes with GST-MS2 at 1.5 µM final concentration at 30 °C for 20 min, after which increasing amounts (0.2, 0.3, 0.7, and 1.3 µM) of hnRNP A1 or UP1 were added, with further incubation for 20 min at 30 °C. Glutathione-agarose beads were added and incubated at 4 °C for 1 hr, followed by washing the beads and elution and extraction of the RNA as described (Mayeda & Krainer, 1999a; Zhang & Krainer, 2007).

2.5.6 RNA Footprinting

Hydroxyl radical and RNase A footprinting experiments were done as described (Clarke, 1999) with 16 nM final concentration of 5’ 32P-labeled RNA, and 1, 2, 4, and 8 pmol (0.05, 0.1, 0.2, and 0.4 µM) of recombinant hnRNP A1. Binding of hnRNP A1 to RNA before incubation with hydroxyl radical or RNase A was done as described in section 2.5.4.
2.6 Acknowledgments

We thank Michele Hastings, Zuo Zhang, and Qingshuo Zhang for generous gifts of recombinant proteins, and Mads Jensen, Olga Anczuków-Camarda, Isabel Aznarez, and Yimin Hua for helpful comments on the manuscript. This work was supported by grant CA13106 from the NCI.

2.7 References


NASIM, F. U., HUTCHISON, S., CORDEAU, M. & CHABOT, B. (2002). High-affinity hnRNP A1 binding sites and duplex-forming inverted repeats have similar effects on 5' splice site selection in support of a common looping out and repression mechanism. Rna 8, 1078-89.


Chapter 3

Consensus Sequences for hnRNP A1
3.1 Abstract

One of the most abundant nuclear proteins is hnRNP A1, which binds to single-stranded RNA and regulates alternative splicing. A robust consensus motif for hnRNP A1 recognition is still lacking, although sequences to which hnRNP A1 binds have been reported, as well as mutations in these sequences that abolish hnRNP A1 binding. Here, we derive an hnRNP A1 consensus motif based on functional in vitro splicing assays. Approximately 200 hexamers that may or may not bind hnRNP A1 were engineered into the 3’ end of exon 2 of a β-globin minigene, and binding assays were carried out using competitive UV cross-linking of 50 randomly picked hexamers that did or did not repress splicing.
3.2 Introduction

Heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) is one of the most abundant nuclear proteins (Dreyfuss et al., 1993; Hanamura et al., 1998) and acts as a splicing factor, binding to cis-acting regulatory elements, such as exonic or intronic splicing silencers (ESS or ISS) to repress splicing in the context of alternative splicing events (Amendt, Si & Stoltzfus, 1995; Burd & Dreyfuss, 1994; Caputi et al., 1994; Chabot et al., 1997; Chew, Baginsky & Eperon, 2000; Del Gatto-Konczak et al., 1999; Expert-Bezancon et al., 2004; Ma et al., 2002; Min, Chan & Black, 1995; Nasim et al., 2002; Paradis et al., 2007; Zheng et al., 2000; Zhu, Mayeda & Krainer, 2001). hnRNP A1 shares features with many other hnRNP proteins, namely two RRMs and an auxiliary domain believed to be responsible for protein-protein, RNA-protein, and single-stranded DNA-protein interactions. hnRNP A1 can also form homophilic interactions in solution and heterophilic interactions with other hnRNPs (Cartegni et al., 1996; Cobianchi et al., 1988; Nadler et al., 1991). hnRNP A1 undergoes cooperative binding to RNA, which is initiated at a high-affinity binding site, or ESS, and spreads along an exon to inhibit splicing (Zhu et al., 2001). hnRNP A1 has antagonistic effects on SR proteins binding to nearby ESEs (Cartegni et al., 1996; Expert-Bezancon et al., 2004; Zhu et al., 2001). The prevalence of hnRNP A1 over SR proteins depends on whether or not hnRNP A1 can displace the bound SR proteins from the ESEs. If hnRNP A1 prevails, there is splicing repression, and if the SR proteins block hnRNP A1 cooperative binding from spreading, splicing can take place (Zhu et al., 2001). Even though a high-affinity binding site consensus motif for hnRNP A1 was discovered through conventional SELEX (Burd & Dreyfuss, 1994), hnRNP A1 can bind to a variety of other sequences, perhaps because all of its domains contribute to binding affinity (Cartegni et al., 1996; Ding et al., 1999). There have been previous attempts to derive a consensus sequence for hnRNP A1 binding (Cartegni et al., 2006; Nielsen et al., 2007); however, the approach employed involved only gathering and analysis of some of the reported hnRNP A1 binding sites. The approach we used also involved an analysis of most of the previously reported hnRNP A1 binding sites, but then used the results of this analysis to construct about 200 beta-globin minigenes for functional assays, to see how each hnRNP A1 binding site represses splicing in S100 extract complemented with SC35. The results obtained from this functional assay can be used to design ribo-oligonucleotides for binding assays.
3.3 Results

3.3.1 Nβ-globin constructs

About 200 hexamers that were randomly generated, some of them are collections of known hexamer from the literatures, of these collections, some are known to bind hnRNP A1 or other hnRNPs while others may not bind hnRNP A1. These hexamers were engineered into the 3’ end of exon 2 of the Nβ-globin construct, as described (Chapter 2) (see Figure 12). Three controls were used: C1, Nβ-globin without the hnRNP A1 binding hexamer; C2, containing a high-affinity hnRNP A1 binding hexamer (UAGGGU) as defined by SELEX (Burd & Dreyfuss, 1994); and C3, containing a binding-defective hexamer mutant (UCGCGG) in which the A at the second position was mutated to C (Chapter 2).

![Diagram of Nβ-globin constructs](image)

*Figure 12. Nβ-globin constructs.*

β-globin minigene transcripts for in vitro splicing assays. The pre-mRNAs comprise 108-nt of exon 1, the 130-nt first intron, and 108-nt of exon 2. NβESS and NβESSm have an additional 6-nt ESS or ESSm, respectively, at the 3’ end. The red bar represents the ESS hexamer; the green bar represents the ESS mutant hexamer (ESSm).
3.3.2 Splicing assay results

All splicing reactions were carried out in a splicing inactive cytoplasmic extract (S100) complemented with a recombinant SR protein (SC35) and recombinant hnRNP A1. For the controls, recombinant hnRNP A1 was omitted (cf. Figures 13 and 14). The splicing reactions were divided into four pools and the results for each pool and pool control are shown in Figures 13 through 20; summaries of the results for each pool are shown in Tables 3 through 6. Percentage relative splicing shown in each table is calculated with respect to the controls (mRNA/pre-mRNA)/(mRNA control/pre-mRNA control)x100%.

**Splicing of β-globin ESS Constructs in S100 + SC35 + hnRNP A1 (Pool 1)**

![Figure 13. Splicing assay results. Pool 1](image)

Splicing of the pre-mRNAs from Figure 12, in which ESS represents 32 different hexamers, in S100 extract complemented with SC35, in the presence of 7.5 pmol of hnRNP A1. Pu represents the insertion of a purine before the high affinity hnRNP A1 binding hexamer. C1 is a Nβ control that lacks an ESS, and C3 is a mutant control with a high-affinity hnRNP A1 binding site. The relative splicing efficiency (mRNA/pre-mRNA)/(mRNA control/pre-mRNA control)x100% is shown on Table 3.
Splicing of β-globin ESS Constructs in S100 + SC35 Control

Figure 14. Splicing assay results. Pool 1 Control
Splicing assays in S100 extract complemented with SC35 in the absence of recombinant hnRNP A1 are the splicing assay controls for Figure 13.

Table 3. Splicing assay results Pool 1.
Relative percentage is calculated as \((\text{mRNA}/\text{pre-mRNA})/(\text{mRNA control}/\text{pre-mRNA control})\times 100\%\).

<table>
<thead>
<tr>
<th>Names</th>
<th>Sequences</th>
<th>% Relative Splicing</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnR0</td>
<td>UAGGGU</td>
<td>0</td>
</tr>
<tr>
<td>C1</td>
<td>UAGGGU</td>
<td>100</td>
</tr>
<tr>
<td>C2</td>
<td>UAGGGU</td>
<td>100</td>
</tr>
<tr>
<td>lnR1</td>
<td>UAGGGU</td>
<td>0</td>
</tr>
<tr>
<td>lnR2</td>
<td>UAGGGU</td>
<td>1.5</td>
</tr>
<tr>
<td>lnR3</td>
<td>UAGGGC</td>
<td>1</td>
</tr>
<tr>
<td>lnR4</td>
<td>UAGGGG</td>
<td>1.5</td>
</tr>
<tr>
<td>lnR5</td>
<td>UAGGAU</td>
<td>3</td>
</tr>
<tr>
<td>lnR6</td>
<td>UAGGAA</td>
<td>3</td>
</tr>
<tr>
<td>lnR7</td>
<td>UAGGAC</td>
<td>3</td>
</tr>
<tr>
<td>lnR8</td>
<td>UAGGAG</td>
<td>10</td>
</tr>
<tr>
<td>lnR9</td>
<td>UAGAGU</td>
<td>3</td>
</tr>
<tr>
<td>lnR10</td>
<td>UAGAGA</td>
<td>10</td>
</tr>
<tr>
<td>lnR11</td>
<td>UAGAGC</td>
<td>10</td>
</tr>
<tr>
<td>lnR12</td>
<td>UAGAGG</td>
<td>33</td>
</tr>
<tr>
<td>lnR13</td>
<td>UGGGGU</td>
<td>60</td>
</tr>
<tr>
<td>lnR14</td>
<td>UGGGGA</td>
<td>60</td>
</tr>
<tr>
<td>lnR15</td>
<td>UGGGGC</td>
<td>50</td>
</tr>
<tr>
<td>lnR16</td>
<td>UGGGGG</td>
<td>75</td>
</tr>
<tr>
<td>lnR17</td>
<td>CAGGGU</td>
<td>7.5</td>
</tr>
<tr>
<td>lnR18</td>
<td>CAGGGA</td>
<td>20</td>
</tr>
<tr>
<td>lnR19</td>
<td>CAGGGC</td>
<td>20</td>
</tr>
<tr>
<td>lnR20</td>
<td>CAGGGG</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Names</th>
<th>Sequences</th>
<th>% Relative Splicing</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnR21</td>
<td>CAGGAU</td>
<td>15</td>
</tr>
<tr>
<td>lnR22</td>
<td>CAGGCA</td>
<td>15</td>
</tr>
<tr>
<td>lnR23</td>
<td>CAGGAC</td>
<td>15</td>
</tr>
<tr>
<td>lnR24</td>
<td>CAGGAG</td>
<td>15</td>
</tr>
<tr>
<td>lnR25</td>
<td>CAGAGU</td>
<td>60</td>
</tr>
<tr>
<td>lnR26</td>
<td>CAGACA</td>
<td>60</td>
</tr>
<tr>
<td>lnR27</td>
<td>CAGAGC</td>
<td>60</td>
</tr>
<tr>
<td>lnR28</td>
<td>CAGAGG</td>
<td>60</td>
</tr>
<tr>
<td>lnR29</td>
<td>CGGGGU</td>
<td>75</td>
</tr>
<tr>
<td>lnR30</td>
<td>CGGGGG</td>
<td>65</td>
</tr>
<tr>
<td>lnR31</td>
<td>CGGGGG</td>
<td>65</td>
</tr>
<tr>
<td>lnR32</td>
<td>CGGGGG</td>
<td>65</td>
</tr>
</tbody>
</table>
Figure 15. Splicing assay results. Pool 2
Splicing of the pre-mRNAs from Figure 12, in which the ESS consists of 44 different hexamers, in S100 extract complemented with SC35, in the presence of 7.5 pmol of hnRNP A1. Relative splicing efficiency is shown on Table 4.
Figure 16. Splicing assay results. Pool 2 Control. The splicing assays in S100 complemented with SC35 in the absence of recombinant hnRNP A1 are the splicing assay controls for Figure 15.
Table 4. Splicing assay results Pool 2.
Relative percentage is calculated as \((\text{mRNA/pre-mRNA})/(\text{mRNA control/pre-mRNA control})\times 100\%\).

<table>
<thead>
<tr>
<th>Names</th>
<th>Sequences</th>
<th>% Relative Splicing</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>UAGGGU</td>
<td>100</td>
</tr>
<tr>
<td>C2</td>
<td>UAGAAA</td>
<td>9</td>
</tr>
<tr>
<td>hnR33</td>
<td>UAGAAG</td>
<td>10</td>
</tr>
<tr>
<td>hnR34</td>
<td>UAGACA</td>
<td>9</td>
</tr>
<tr>
<td>hnR35</td>
<td>UAGAAU</td>
<td>6</td>
</tr>
<tr>
<td>hnR36</td>
<td>UAGAUC</td>
<td>9</td>
</tr>
<tr>
<td>hnR37</td>
<td>UAGAC</td>
<td>9</td>
</tr>
<tr>
<td>hnR38</td>
<td>UAGACU</td>
<td>6</td>
</tr>
<tr>
<td>hnR39</td>
<td>UAGAU</td>
<td>4</td>
</tr>
<tr>
<td>hnR40</td>
<td>UAGAG</td>
<td>3</td>
</tr>
<tr>
<td>hnR41</td>
<td>UAGAUU</td>
<td>0</td>
</tr>
<tr>
<td>hnR42</td>
<td>UAGCGA</td>
<td>3</td>
</tr>
<tr>
<td>hnR43</td>
<td>UAGCGC</td>
<td>2</td>
</tr>
<tr>
<td>hnR44</td>
<td>UAGCGG</td>
<td>7</td>
</tr>
<tr>
<td>hnR45</td>
<td>UAGCGU</td>
<td>2</td>
</tr>
<tr>
<td>hnR46</td>
<td>UAGCCA</td>
<td>2</td>
</tr>
<tr>
<td>hnR47</td>
<td>UAGGCC</td>
<td>2</td>
</tr>
<tr>
<td>hnR48</td>
<td>UAGGCA</td>
<td>2</td>
</tr>
<tr>
<td>hnR49</td>
<td>UAGGCC</td>
<td>2</td>
</tr>
<tr>
<td>hnR50</td>
<td>UAGGCC</td>
<td>2</td>
</tr>
<tr>
<td>hnR51</td>
<td>UAGGCG</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Names</th>
<th>Sequences</th>
<th>% Relative Splicing</th>
</tr>
</thead>
<tbody>
<tr>
<td>hnR52</td>
<td>UAGGCU</td>
<td>0</td>
</tr>
<tr>
<td>hnR53</td>
<td>UAGGUA</td>
<td>2</td>
</tr>
<tr>
<td>hnR54</td>
<td>UAGGUC</td>
<td>0</td>
</tr>
<tr>
<td>hnR55</td>
<td>UAGGUG</td>
<td>2</td>
</tr>
<tr>
<td>hnR56</td>
<td>UAGGUU</td>
<td>0</td>
</tr>
<tr>
<td>hnR57</td>
<td>UAGUAA</td>
<td>2</td>
</tr>
<tr>
<td>hnR58</td>
<td>UAGUAC</td>
<td>2</td>
</tr>
<tr>
<td>hnR59</td>
<td>UAGUAG</td>
<td>2</td>
</tr>
<tr>
<td>hnR60</td>
<td>UAGUAU</td>
<td>2</td>
</tr>
<tr>
<td>hnR61</td>
<td>UAGUGA</td>
<td>2</td>
</tr>
<tr>
<td>hnR62</td>
<td>UAGUGC</td>
<td>2</td>
</tr>
<tr>
<td>hnR63</td>
<td>UAGUGG</td>
<td>2</td>
</tr>
<tr>
<td>hnR64</td>
<td>UAGUGU</td>
<td>2</td>
</tr>
<tr>
<td>hnR65</td>
<td>UAGUUA</td>
<td>9</td>
</tr>
<tr>
<td>hnR66</td>
<td>UAGUUC</td>
<td>50</td>
</tr>
<tr>
<td>hnR67</td>
<td>UAGUUG</td>
<td>9</td>
</tr>
<tr>
<td>hnR68</td>
<td>UAGUUU</td>
<td>9</td>
</tr>
<tr>
<td>hnR69</td>
<td>UGGAGA</td>
<td>110</td>
</tr>
<tr>
<td>hnR70</td>
<td>UGGAGC</td>
<td>100</td>
</tr>
<tr>
<td>hnR71</td>
<td>UGGAGG</td>
<td>100</td>
</tr>
<tr>
<td>hnR72</td>
<td>UGGAGU</td>
<td>100</td>
</tr>
<tr>
<td>hnR73</td>
<td>UGGGAA</td>
<td>100</td>
</tr>
<tr>
<td>hnR74</td>
<td>UGGGAC</td>
<td>100</td>
</tr>
<tr>
<td>hnR75</td>
<td>UGGGAG</td>
<td>100</td>
</tr>
<tr>
<td>hnR76</td>
<td>UGGGAU</td>
<td>100</td>
</tr>
</tbody>
</table>
Figure 17. Splicing assay results. Pool 3
Splicing of the pre-mRNAs from Figure 12, in which the ESS consists of 52 different hexamers, in S100 extract complemented with SC35, in the presence of 7.5 pmol of hnRNP A1. Relative splicing efficiency is shown on Table 5.
**Figure 18.** Splicing assay results. Pool 3 Control.
The splicing assays in S100 complemented with SC35 in the absence of recombinant hnRNP A1 are the splicing assay controls for Figure 17.
Table 5. Splicing assay results Pool 3.
Relative percentage is calculated as \((\text{mRNA/pre-mRNA})/(\text{mRNA control/pre-mRNA control})\times100\%\).

<table>
<thead>
<tr>
<th>Names</th>
<th>Sequences</th>
<th>% Relative Splicing</th>
</tr>
</thead>
<tbody>
<tr>
<td>hnR77</td>
<td>UGGGUA</td>
<td>60</td>
</tr>
<tr>
<td>hnR78</td>
<td>UGGGUC</td>
<td>100</td>
</tr>
<tr>
<td>hnR79</td>
<td>UGGGUG</td>
<td>95</td>
</tr>
<tr>
<td>hnR80</td>
<td>UGGGUU</td>
<td>70</td>
</tr>
<tr>
<td>hnR81</td>
<td>UGGUAA</td>
<td>70</td>
</tr>
<tr>
<td>hnR82</td>
<td>UGGUAC</td>
<td>75</td>
</tr>
<tr>
<td>hnR83</td>
<td>UGGUAG</td>
<td>3</td>
</tr>
<tr>
<td>hnR84</td>
<td>UGGUAU</td>
<td>60</td>
</tr>
<tr>
<td>hnR85</td>
<td>UGGUGA</td>
<td>100</td>
</tr>
<tr>
<td>hnR86</td>
<td>UGGUGC</td>
<td>100</td>
</tr>
<tr>
<td>hnR87</td>
<td>UGGUGG</td>
<td>70</td>
</tr>
<tr>
<td>hnR88</td>
<td>UGGUGU</td>
<td>100</td>
</tr>
<tr>
<td>hnR89</td>
<td>UGGUUA</td>
<td>85</td>
</tr>
<tr>
<td>hnR90</td>
<td>UGGUUUC</td>
<td>90</td>
</tr>
<tr>
<td>hnR91</td>
<td>UGGUUG</td>
<td>90</td>
</tr>
<tr>
<td>hnR92</td>
<td>UGGUUU</td>
<td>70</td>
</tr>
<tr>
<td>hnR93</td>
<td>CAGAAA</td>
<td>70</td>
</tr>
<tr>
<td>hnR94</td>
<td>CAGAAC</td>
<td>80</td>
</tr>
<tr>
<td>hnR95</td>
<td>CAGAAC</td>
<td>95</td>
</tr>
<tr>
<td>hnR96</td>
<td>CAGAUA</td>
<td>20</td>
</tr>
<tr>
<td>hnR97</td>
<td>CAGACA</td>
<td>50</td>
</tr>
<tr>
<td>hnR98</td>
<td>CAGACCC</td>
<td>80</td>
</tr>
<tr>
<td>hnR99</td>
<td>CAGACG</td>
<td>95</td>
</tr>
<tr>
<td>hnR100</td>
<td>CAGACU</td>
<td>43</td>
</tr>
<tr>
<td>hnR101</td>
<td>CAGAUA</td>
<td>30</td>
</tr>
<tr>
<td>hnR102</td>
<td>CAGAUC</td>
<td>70</td>
</tr>
<tr>
<td>Names</td>
<td>Sequences</td>
<td>% Relative Splicing</td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>---------------------</td>
</tr>
<tr>
<td>hnR103</td>
<td>CAGAUG</td>
<td>85</td>
</tr>
<tr>
<td>hnR104</td>
<td>CAGAUA</td>
<td>50</td>
</tr>
<tr>
<td>hnR105</td>
<td>CAGCGA</td>
<td>70</td>
</tr>
<tr>
<td>hnR106</td>
<td>CAGCGC</td>
<td>70</td>
</tr>
<tr>
<td>hnR107</td>
<td>CAGCGG</td>
<td>70</td>
</tr>
<tr>
<td>hnR108</td>
<td>CAGCGU</td>
<td>70</td>
</tr>
<tr>
<td>hnR109</td>
<td>CAGGAA</td>
<td>70</td>
</tr>
<tr>
<td>hnR110</td>
<td>CAGGCC</td>
<td>100</td>
</tr>
<tr>
<td>hnR111</td>
<td>CAGGCG</td>
<td>100</td>
</tr>
<tr>
<td>hnR112</td>
<td>CAGGCU</td>
<td>30</td>
</tr>
<tr>
<td>hnR113</td>
<td>CGGGUA</td>
<td>100</td>
</tr>
<tr>
<td>hnR114</td>
<td>CGGGUC</td>
<td>100</td>
</tr>
<tr>
<td>hnR115</td>
<td>CGGGUG</td>
<td>100</td>
</tr>
<tr>
<td>hnR116</td>
<td>CGGGUU</td>
<td>100</td>
</tr>
<tr>
<td>hnR117</td>
<td>CGGUAA</td>
<td>100</td>
</tr>
<tr>
<td>hnR118</td>
<td>CGGUAC</td>
<td>100</td>
</tr>
<tr>
<td>hnR119</td>
<td>CGGUAG</td>
<td>100</td>
</tr>
<tr>
<td>hnR120</td>
<td>CGGUUA</td>
<td>100</td>
</tr>
<tr>
<td>hnR121</td>
<td>CGGUGA</td>
<td>100</td>
</tr>
<tr>
<td>hnR122</td>
<td>CGGUGC</td>
<td>100</td>
</tr>
<tr>
<td>hnR123</td>
<td>CGGUGG</td>
<td>100</td>
</tr>
<tr>
<td>hnR124</td>
<td>CGGUGU</td>
<td>100</td>
</tr>
<tr>
<td>hnR125</td>
<td>CGGUAU</td>
<td>100</td>
</tr>
<tr>
<td>hnR126</td>
<td>CGGUUC</td>
<td>100</td>
</tr>
<tr>
<td>hnR127</td>
<td>CGGUUG</td>
<td>100</td>
</tr>
<tr>
<td>hnR128</td>
<td>CGGUUU</td>
<td>100</td>
</tr>
</tbody>
</table>
Figure 19. Splicing assay results. Pool 4
Splicing of the pre-mRNAs from Figure 12, in which the ESS consists of 32 different hexamers, in S100 extract complemented with SC35, in the presence of 7.5 pmol of hnRNP A1. C1 (Nβ) lacks the high-affinity hnRNP A1 binding site. The ESS control, C2, has a high-affinity hnRNP A1 binding site as the ESS. C3 is a mutant control with a known high-affinity hnRNP A1 binding site. Relative splicing efficiency is shown on Table 6.

Figure 20. Splicing assay results. Pool 4 Control
The splicing assays in S100 complemented with SC35 in the absence of recombinant hnRNP A1 are the splicing assay controls for Figure 19.
Relative percentage is calculated as \((\text{mRNA/pre-mRNA})/(\text{mRNA control/pre-mRNA control}) \times 100\%\).

<table>
<thead>
<tr>
<th>Names</th>
<th>Sequences</th>
<th>% Relative Splicing</th>
<th>Names</th>
<th>Sequences</th>
<th>% Relative Splicing</th>
</tr>
</thead>
<tbody>
<tr>
<td>hnR129</td>
<td>UAGAU A</td>
<td>7</td>
<td>hnR145</td>
<td>CAGUGA</td>
<td>50</td>
</tr>
<tr>
<td>hnR130</td>
<td>UAGAUC</td>
<td>7</td>
<td>hnR146</td>
<td>CAGUGC</td>
<td>50</td>
</tr>
<tr>
<td>hnR131</td>
<td>UAGAUG</td>
<td>7</td>
<td>hnR147</td>
<td>CAGUGG</td>
<td>0</td>
</tr>
<tr>
<td>hnR132</td>
<td>UAGAUU</td>
<td>7</td>
<td>hnR148</td>
<td>CAGUGU</td>
<td>30</td>
</tr>
<tr>
<td>hnR133</td>
<td>UAGCGA</td>
<td>20</td>
<td>hnR149</td>
<td>CAGUUA</td>
<td>20</td>
</tr>
<tr>
<td>hnR134</td>
<td>UAGCGC</td>
<td>20</td>
<td>hnR150</td>
<td>CAGUUC</td>
<td>30</td>
</tr>
<tr>
<td>hnR135</td>
<td>UAGCGG</td>
<td>20</td>
<td>hnR151</td>
<td>CAGUUG</td>
<td>25</td>
</tr>
<tr>
<td>hnR136</td>
<td>UAGCGU</td>
<td>10</td>
<td>hnR152</td>
<td>CAGUUU</td>
<td>25</td>
</tr>
<tr>
<td>hnR137</td>
<td>CAGGUA</td>
<td>15</td>
<td>hnR153</td>
<td>UAGGUU</td>
<td>10</td>
</tr>
<tr>
<td>hnR138</td>
<td>CAGGUC</td>
<td>15</td>
<td>hnR154</td>
<td>UAUGGU</td>
<td>15</td>
</tr>
<tr>
<td>hnR139</td>
<td>CAGGUG</td>
<td>20</td>
<td>hnR155</td>
<td>UAAAAA</td>
<td>10</td>
</tr>
<tr>
<td>hnR140</td>
<td>CAGGUU</td>
<td>10</td>
<td>hnR156</td>
<td>UAAAGU</td>
<td>10</td>
</tr>
<tr>
<td>hnR141</td>
<td>CAGUAA</td>
<td>15</td>
<td>hnR157</td>
<td>CAAGGU</td>
<td>50</td>
</tr>
<tr>
<td>hnR142</td>
<td>CAGUAC</td>
<td>20</td>
<td>hnR158</td>
<td>CAUGGU</td>
<td>50</td>
</tr>
<tr>
<td>hnR143</td>
<td>CAGUAG</td>
<td>20</td>
<td>hnR159</td>
<td>CAAAAA</td>
<td>50</td>
</tr>
<tr>
<td>hnR144</td>
<td>CAGUAU</td>
<td>20</td>
<td>hnR160</td>
<td>CAAAGU</td>
<td>50</td>
</tr>
</tbody>
</table>

3.3.3 Matrix and the consensus sequences for hnRNP A1

The position weight matrix and the consensus sequences generated as shown in Figure 21 are based on the splicing assays summarized in Tables 3 through 6, and have yet to be confirmed by binding experiments. To construct a position weight matrix, a frequency matrix \(f_i(a)\) was first calculated from the alignment \((i\text{ is the position of nucleotide } a)\). Given a background frequency for the set of sequences, \(p(a)\), the scoring matrix is defined by the following formula:

\[
s_i(a) = \log_2 \left( \frac{f_i(a) + \epsilon p(a)}{p(a)(1 + \epsilon)} \right)
\]

where \(i = (1, 2, ..., L)\), \(a = (A, C, G, U)\), and \(\epsilon = 0.5\) is the Bayesian prior parameter (Lawrence et al., 1993; Liu, Zhang & Krainer, 1998). A motif score is equal to the sum of the scores at each position. Motifs may be ranked by their scores. The top three scores in each sequence using different scoring matrices were calculated as described (Liu et al., 1998). The sequence scores were consistent semiquantitatively with the
percentage splicing inhibition data when the sequence-scores for hnRNP A1 were defined as described (Liu et al., 1998).

Matrix and the consensus sequences for hnRNP A1

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-1.585</td>
<td>1.5</td>
<td>-1.24</td>
<td>0.082</td>
<td>-0.05</td>
<td>-0.05</td>
</tr>
<tr>
<td>C</td>
<td>0.1594</td>
<td>-2</td>
<td>-1.58</td>
<td>-0.78</td>
<td>-0.56</td>
<td>-0.21</td>
</tr>
<tr>
<td>G</td>
<td>-1.585</td>
<td>-1</td>
<td>1.5</td>
<td>0.4</td>
<td>0.295</td>
<td>-0.03</td>
</tr>
<tr>
<td>T</td>
<td>1.2216</td>
<td>-2</td>
<td>-1.45</td>
<td>0.047</td>
<td>0.18</td>
<td>0.262</td>
</tr>
</tbody>
</table>

Figure 21. Matrix representing hnRNP A1 binding consensus sequences. This position weight matrix and the hnRNP A1 consensus sequences generated are based on the results presented in Tables 3 to 6.

3.4 Discussion

Even though we generated the consensus sequences based on the splicing assays alone, we feel that this experiment is not complete until we confirm and extend our results with an appropriate binding assay, which is currently under way.
3.5 Methods and Materials

3.5.1 Transcripts

All the pre-mRNA transcripts for in vitro splicing were capped, and \(^{32}\)P labeled by in vitro transcription in the presence of [\(\alpha\)-\(^{32}\)P]-UTP from PCR templates with a T7 phage promoter (Mayeda & Krainer, 1999). The PCR primers for all the transcription templates of all constructs used for the in vitro splicing experiments are available upon request. The template for PCR to generate all the constructs was the the linearized product of \(Hind\) III restriction digest of plasmid pSP64-H\(\beta\)\(\Delta\)6 (Krainer et al., 1984; Mayeda et al., 1999).

3.5.2 Recombinant Proteins

Untagged human hnRNP A1 was expressed in E. coli and purified as described (Mayeda et al., 1994). Purified human SC35 expressed in baculovirus was a gift from Michelle Hastings.

3.5.3 In vitro splicing Assays

S100 extract from HeLa cells was prepared as described (Caputi et al., 1999). In vitro splicing reactions were carried out in a final volume of 12.5 µl with 15 fmol (1.15 nM) of \(^{32}\)P-labeled, \(^7\)CH\(_3\)-GpppG-capped T7 RNA transcripts, 35% (v/v) S100 extract, with a final concentration of 0.4
µM SC35, in the presence or absence of hnRNP A1 at a final concentration of 0.6 µM.

3.6 References


Chapter 4

Summary
4.1 Summary

The previous two chapters focused on hnRNP A1 and the mechanism by which it represses splicing. In chapter two, we showed that hnRNP A1 binding to RNA and splicing repression can occur on an unstructured RNA. Moreover, hnRNP A1 can effectively unwind RNA hairpins upon binding. We also showed that hnRNP A1 can spread in a 5’ to 3’ direction, although when initial binding takes place in the middle of an RNA, spreading proceeds preferentially in a 3’ to 5’ direction. Finally, when two distant high-affinity sites are present, they can facilitate cooperative spreading of hnRNP A1 between the two sites. In Chapter 3, we derive an hnRNP A1 consensus motif based on functional in vitro splicing assays. There are other hnRNP proteins that also have inhibitory effects on splicing. One of them is hnRNP I/PTB (polypyrimidine-tract binding protein) whose mechanism of inhibition has been extensively studied (Bonderoff, Larey & Lloyd, 2008; Grover, Ray & Das, 2008; Kuwahata et al., 2008; Lewis, Gagnon & Mowry, 2008; Paradis et al., 2007; Radzimanowski et al., 2008; Sawicka et al., 2008; Wang et al., 2008). However, it is unknown whether PTB can undergo cooperative binding or cooperative interactions with hnRNP A1 to inhibit splicing (Bolanos-Garcia, 2005; Borg & Margolis, 1998; Fred, Tillmar & Welsh, 2006; Margolis, 1996; Sawicka et al., 2008; Shoelson, 1997). At the beginning of my study, when I was looking for sequences other than the hexamer at the 3’ end of exon 3 of HIV-1 tat 23 that might bind hnRNP A1 better, one of the sequences that I tested was a polypyrimidine tract; this sequence inhibited splicing of tat 23 better than the tat 23 natural hnRNP A1 binding site but it did it in the absence of hnRNP A1 recombinant protein. The manner of inhibition of splicing by PTB on this tat 23 mutant was very similar to that of hnRNP A1, so that it might be possible that the mechanism of inhibition of splicing by PTB or any other inhibitory hnRNP proteins could also be cooperative spreading like hnRNP A1. It is also possible that some of these hnRNPs can act cooperatively in conjunction with hnRNP A1 to inhibit splicing. If I had more time, I would like to have categorized the interactions between hnRNP A1 and the other hnRNPs. I would also have liked to study some of these hnRNPs in cooperative binding studies similar to the one described in Chapter 2.
4.2 References


Bibliography


Regulation of alternative splicing in vivo by overexpression of antagonistic C and origins of U12 splicing.


Evolutionary fates and origins of U12-type introns. Mol Cell 2, 773-85.


hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. Embo J 18, 4060-7.

hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. Embo J 18, 4060-7.


HJ the human telomeric DNA sequence d(TTAGGG)n. 

proteins that bind the pre splicing by a novel repeated hexanucleotide element. 

Mol Cell H in Drosophila. 

Control of doublesex alternative splicing by transformer and transformer enhancers. 

J Biol Chem characterization of nucleic acid helix RNA polyadenylation of dsx pre and characterization of a kinase specific for the serine-arginine-rich proteins. 

Evidence by sequence comparison of a new family of mammalian protein kinases. 

Characterization by cDNA cloning of a kinase specific for the serine-arginine-rich proteins. 

hnRNP A1 is required for processing of miR-18a. 


Rna 4, 430-44. 

Evidence by sequence comparison of a new family of mammalian protein kinases. 

J Mol Biol 244, 665-72. 


Cell 65, 579-86. 

J Cell Biol 78, 663-74. 

J Biol Chem 251, 2124-32. 

The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. 


Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors. 

Rna 4, 430-44. 

Characterization by cDNA cloning of two new human protein kinases. 

J Mol Biol 244, 665-72. 

Pre-mRNA splicing in the new millennium. 

Sex-specific splicing and polyadenylation of dsx pre-mRNA requires a sequence that binds specifically to tra-2 protein in vitro. 

Cell 65, 579-86. 

Heterogeneous nuclear RNA-protein fibers in chromatin-depleted nuclei. 

J Cell Biol 78, 663-74. 

Purification and physical characterization of nucleic acid helix-unwinding proteins from calf thymus. 

J Biol Chem 251, 2124-32. 

The function of multisite splicing enhancers. 

Mol Cell 1, 449-55. 

Widespread and subtle: alternative splicing at short-distance tandem sites. 

Trends Genet 24, 246-55. 

Control of doublesex alternative splicing by transformer and transformer-2 in Drosophila. 

Science 252, 833-6. 

SRprises along a messenger's journey. 

Mol Cell 17, 613-5. 

Regulation of alternative pre-mRNA splicing by a novel repeated hexanucleotide element. 

Genes Dev 8, 1561-74. 

Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) and the human telomeric DNA sequence d(TTAGGG)n. 

Mol Cell Biol 13, 4301-10. 

U1 snRNP-ASF/SF2 interaction and 5' splice site recognition: characterization of required elements. 


152. PARADIS, C., CLOUTIER, P., SHKRETA, L., TOUTANT, J., KLARSKOV, K.
& CHABOT, B. (2007). hnRNP I/PTB can antagonize the splicing repressor
cancer therapy: chronic myelogenous leukemia shows the way. *Clin
Cancer Res* **13**, 1638-42.
(2007a). Rapid, transcript-specific changes in splicing in response to
(2007b). Transcript specificity in yeast pre-mRNA splicing revealed by
156. POLLARD, A. J., SPAREY, C., ROBSON, S. C., KRAINER, A. R. &
EUROPE-FINNER, G. N. (2000). Spatio-temporal expression of the trans-
acting splicing factors SF2/ASF and heterogeneous ribonuclear proteins
A1/A1B in the myometrium of the pregnant human uterus: a molecular
mechanism for regulating regional protein isoform expression in vivo. *J
dNA strands mediated by purified mammalian heterogeneous nuclear
ribonucleoprotein A1 protein: implications for a mechanism for rapid
158. PONTIUS, B. W. & BERG, P. (1992). Rapid assembly and
disassembly of complementary DNA strands through an equilibrium
13815-8.
protein kinase Clk/Sty directly modulates SR protein activity: both hyper-
RNA recognition motif identified within a defined U1 RNA binding domain
162. RADZIMANOWSKI, J., RAVAUD, S., SCHLESINGER, S., KOCH, J.,
differences between authentic and cryptic 5' splice sites. *Nucleic Acids
Determinants of the inherent strength of human 5' splice sites. *Rna* **11**,
683-98.


