universität wien

DISSERTATION

Titel der Dissertation

Role of RNA editing in RNA splicing and in nuclear export of microRNA precursors

angestrebter akademischer Grad

Doktor/in der Naturwissenschaften (Dr. rer.nat.)

Verfasserin / Verfasser:	Draško Boko
Matrikel-Nummer:	0547500
Dissertationsgebiet (lt. Studienblatt):	A 091 490
Betreuerin / Betreuer:	Ao. Univ.-Prof. Dr. Michael F. Jantsch

Wien, im August 2009

ABSTRACT

The double-stranded RNA binding domain (dsRBD) is a common motif found in proteins involved in RNA processing, localisation or interference. RNA recognition by these proteins depends on the secondary structure and not on the sequence.

Adenosine deaminases that act on RNA (ADARs) are dsRBD-containing proteins which modify target adenosines into inosines in double-stranded RNA substrates, frequently in vicinity of 5´ splice sites, having a profound impact on the subsequent RNA processing events, including RNA splicing and stability. Accurate splicing of messenger RNAs coding for the subunit B of the glutamate receptor, for example, relies upon ADAR-mediated RNA editing at several positions, including the R/G site located in exon 13, just two nucleotides upstream of the exon-intron border. The exact macromolecular interaction affected by deamination at this site, however, has not been identified. We postulated that inosine in this critical position influences U1 snRNA base-pairing to this imperfect 5´ splice site. This hypothesis was tested by employing compensatory U1 snRNA mutants. Furthermore, the sequence surrounding the edited R/G site strongly resembles a splicing silencer consensus bound by splicing repressor proteins, like hnRNP A1 or hnRNP H. Therefore, the interaction between R/G site and these proteins was also investigated.

MicroRNA precursors represent another prominent class of ADAR substrates whose processing is substantially altered upon RNA editing, as they become sensitive to cleavage by Tudor-SN, a cytoplasmic ribonuclease. Since ADAR1 is a dsRNA-binding shuttling protein that interacts with Exportin5, a main export receptor for miRNA precursors, a co-immunoprecipitation method was exerted to detect the interaction of ADAR1 and edited miRNA precursors in both nucleus and cytoplasm.

In eukaryotic genomes, short interspersed elements (SINEs) are transposable regions mobilised through an RNA intermediate that interacts with dsRBD-containing proteins crucial for transcription and translation control. In this part of the study, it is being shown that plant SB1 SINE RNA interacts strongly with imperfect dsRNA-binding protein HYL1, a key factor in the microRNA and trans-acting small interfering RNA (tasiRNA) production in plants, but not with perfect dsRNA-specific protein DRB4 *in vitro*. The binding site maps to stem-loop structure highly reminiscent of miRNA precursors, suggesting how SINE RNAs could regulate different RNAi pathways.

ZUSAMMENFASSUNG

Doppelstrang RNA-bindende Domänen (dsRBDs) sind häufige Motive in Proteinen die in RNA Prozessierung, Lokalisierung oder RNA Interferenz involviert sind. Für RNA Bindung spielt dabei die Sekundärstruktur der RNA und nicht ihre Sequenz die entscheidende Rolle.

„Adenosine deaminases that act on RNA" (ADARs) sind dsRBD-beinhaltende Proteine, deren Aufgabe die Deaminierung von Adenosin zu Inosin in spezifischen doppelsträngigen RNA Substraten ist. Die modifizierten Basen befinden sich häufig nahe an 5´ Spleißstellen und haben infolgedessen einen breiten Einfluß auf RNA Spleißen und Stabilität. Korrektes Spleißen der mRNA der Untereinheit B des Glutamatrezeptors benötigt zum Beispiel RNA Editierung durch ADAR an mehreren Stellen, unter anderem auch an der R/G Position in Exon 13, die nur 2 Nukleotide von der 5´ Spleißstelle liegt. Der genaue Mechanismus, wie Deaminierung an der R/G Position Spleißen beeinflußt, ist jedoch unklar. Wir vermuteten, dass Inosine an dieser kritischen Position die Interaktion der 5´ Spleißstelle mit U1 snRNA ändert. Diese Hypothese wurde experimentell durch eine kompensierende U1 snRNA Mutante getestet. Weiters weist die Umgebung der R/G Stelle Ähnlichkeiten mit spleißhemmenden Sequenzen auf, an denen reprimierende Proteine wie hnRNP A1 und hnRNP H binden. Deswegen wurde auch die Interaktion dieser Faktoren mit der R/G Stelle untersucht.

Die Vorläufer der microRNAs („microRNA precursors") stellen eine weitere bedeutende Klasse von ADAR Substraten dar. Die Prozessierung dieser Vorläufer ist ebenfalls stark durch RNA Editierung geändert. Sie werden, zum Beispiel, suszeptibel für den Abbau durch die cytoplasmatische Ribonuklease Tudor-SN. Da ADAR1 ein dsRNA-bindendes Protein ist, das zwischen Nukleus und Cytoplasma wandert und mit Exportin5, dem Hauptexportrezeptor der microRNA Vorläufer interagiert, wurden co-immunpräzipitierende Methoden angewandt, um die Interaktion von ADAR1 und editierten microRNA Vorläufer sowohl im Nukleus als auch im Cytoplasma zu untersuchen.

In eukaryotischen Genomen findet man „Short Interspersed Elements" (SINEs) als transposable Elemente, die durch ein RNA Intermediat mobilisiert werden. Dieses RNA Intermediat interagiert mit dsRBD-beinhaltenden Proteinen, die wichtig für Transkription und Translation sind. Diesbezüglich wurde gezeigt, dass pflanzliches

SB1 SINE RNA mit dem an partielle doppelsträngige RNA bindenden Protein HYL1, jedoch nicht mit DRB4, das für die perfekt doppelsträngige RNA Struktur spezifisch ist, *in vitro* interagiert. HYL1 ist ein wichtiger Faktor der microRNA und tasiRNA Produktion in Pflanzen. Die Bindestelle für HYL1 an SINE RNA befindet sich an einer „stem-loop“ Struktur, die stark an microRNA Vorläufer erinnert. Dies lässt Schlüsse zu, wie SINE RNAs RNAi Wege bestimmen könnten.

TABLE OF CONTENTS

1 INTRODUCTION 1

Part I: ADAR-mediated RNA editing 2

1.1 Eukaryotic RNA processing 2

1.2 RNA modification 3

1.3 RNA editing 4

1.3.1 Insertion and deletion RNA editing 4

1.3.2 Ribonucleotide deamination 5

1.3.2.1 Cytidine to uridine (C-to-U) RNA editing 6

1.3.2.2 Adenosine to inosine (A-to-I) RNA editing 7

1.4 ADARs (adenosine deaminases that act on RNA) 9

1.4.1 First reports on ADARs 9

1.4.2 Members of the ADAR family 10

1.4.3 Domain architecture of ADARs 12

1.4.4 Substrate recognition by ADARs and the mechanism of adenosine deamination 13

1.4.5 ADAR substrates 15

1.4.5.1 ADAR editing in the central nervous system 15

1.4.5.2 Editing of viral dsRNA 18

1.4.5.3 Non-coding ADAR substrates 19

1.4.6 Biological importance of ADARs 22

1.5 RNA splicing 25

1.5.1 Sequence requirements in RNA splicing 25

1.5.2 Chemistry of RNA splicing 25

1.5.3 Spliceosome assembly 27

1.5.4 Regulation of RNA splicing 29

1.5.5 Connection between RNA editing and RNA splicing 31

Part II: SINE RNAs in RNA interference 33

1.6 SINE RNA interaction with dsRBPs in plants 33

1.6.1 SINE RNAs 33

1.6.2 Plant dsRBPs involved in RNA interference 34

2 SPECIFIC AIMS 36

3 MATERIALS AND METHODS 38

3.1 *In vivo* splicing assay 39

3.1.1 R/G site reporter constructs 39

3.1.2 U1 snRNA expression vectors 39

3.1.3 Primer extension assay 41

3.1.4 Cell culture and transfection 41

3.1.5 FACS analysis 42

3.2 Co-immunoprecipitation and related procedures 42

3.2.1 UV crosslinking 42

3.2.2 Subcellular fractionation 42

3.2.3 Co-immunoprecipitation 43

3.2.4 SDS polyacrylamide gel electrophoresis 43

3.2.5 Western blotting 43

3.2.6 RNA isolation and RT-PCR 44

3.3 Nuclease protection assay 45

4 RESULTS 47

4.1 Molecular mechanism of altered pre-mRNA splicing upon RNA editing . 48

4.1.1 Testing the imperfect 5´splice site hypothesis 49

4.1.1.1 *In* vivo splicing assay 49

4.1.1.2 Reduction of splicing in pre-edited R/G site construct 51

4.1.1.3 Overexpression of U1 snRNA mutants 52

4.1.1.4 Co-transfection experiments 54

4.1.1.5 FACS analysis 55

4.1.2 Interaction of R/G site pre-mRNAs with hnRNPs 61

4.2 ADAR1-mediated export of miRNA precursors 65

4.2.1 ADAR1p150 and miR-22 precursor interact in both nucleus and cytoplasm 66

4.3 SINE RNA interacts with HYL1, key member of RNAi pathway in plants . 70

4.3.1 SINE RNA interacts with a subset of double-stranded RNA binding proteins 71

4.3.2 HYL1 induces conformational changes upon binding to RNA 74

5 DISCUSSION 79

5.1 Compensatory U1 snRNA partially restores the splicing efficiency of pre-edited R/G site construct 80

5.1.1 U1 snRNA interaction with 5´ splice site is disrupted upon RNA editing at R/G site 81

5.1.2 Role for hnRNPs in regulation of splicing through RNA editing 83

5.2 ADAR1p150 remains associated with edited miRNA precursor in the cytoplasm 84

5.3 SINE RNAs bind only a particular subset of dsRBPs 87

5.3.1 SINE RNA Can Bind HYL1 but not DRB4 87

5.3.2 Chaperoning-Like Activity of HYL1 88

5.3.3 Possible evolutionary consequence of the SINE RNA/dsRBP interaction 88

6 REFERENCES 90

ACKNOWLEDGEMENTS 113

CURRICULUM VITAE 114

1 INTRODUCTION

Part I: ADAR-mediated RNA editing

1.1 Eukaryotic RNA processing

Contrary to mRNAs in prokaryotes, whose 5’ ends are already being decoded by the ribosomes while their 3’ ends are still being transcribed, the majority of eukaryotic RNAs need to undergo extensive modification before they are released from the nucleus to the cytoplasm as mature transcripts.

The eukaryotic pre-mRNA molecule is subjected to three main modifications which include 5' capping, 3' polyadenylation, and RNA splicing.

5´ capping of the pre-mRNA involves the addition of 7-methylguanosine (m7G) to the 5' end. First, the terminal 5' phosphate is removed leaving the diphosphate 5' end which, in turn, attacks the α phosphorus atom of a GTP molecule in order to add the guanosine. The guanine ring gets methylated at the N-7 position to form a cap structure. Further methylations may occur at the 2' OH ribose groups of the downstream nucleotides, giving rise to more complex cap structures. The cap protects the 5' end of the primary RNA transcript from cleavage by 5´ exonucleases.

The pre-mRNA processing at the 3' end involves cleavage of its 3' end and subsequent addition of approximately 200 adenosines to form a poly-A tail. Cleavage (5'-CA-3') and polyadenylation (5'- AAUAAA-3') signal sequences direct the cleavage and adenylation reactions. Poly(A) polymerase then adds about 200 adenosine units to the new 3' end of the RNA molecule using ATP as a precursor. As the poly(A) tail is synthesised, it binds multiple copies of poly(A) binding protein, which protects the 3' end from ribonuclease digestion.

RNA splicing as the most remarkable of these alterations can be found in almost every eukaryotic pre-mRNA, which contains coding (exons) and non-coding (introns) regions. Although every intron is flanked by short sequences showing the splicing machinery where to excise the non-coding region, not every splice site is efficiently used. Consequently, mRNAs transcribed from the same gene may vary in length and number of their fused exons. Thus, alternative splicing can provide the cell with a huge set of different proteins derived from a single gene. In fact, this mechanism is presumably one of the most important reasons why there are so many different proteins in contrast to the comparably low number of genes, especially in mammals.

1.2 RNA modification

Chemical and structural diversity and extent of RNA posttranscriptional modification is extraordinary, with over 100 different nucleosides for which chemical structures have been assigned presently known in all types of RNA. The largest number with the greatest structural diversity are found in tRNA, followed by rRNA, mRNA and other RNA species (snRNA). Most modification sites and their positions are conserved, but the differences across distant phylogenetic domains are still reported.

The mechanism of pre-rRNA modification raises special interest because it is guided by the partially complementary small nucleolar RNAs (snoRNAs). The primary rRNA transcript in eukaryotes is 7.5 kb RNA that gets specifically methylated at over 100 sites in higher vertebrates: 80 % of these modifications are 2´-O-ribose methylations, while the rest occurs on either adenine or guanine nitrogenous bases. Additionally, up to 100 uridines in human pre-rRNA get converted into pseudouridines (Ψ). Both of these processes are directed by snoRNAs. These 100-200 nucleotides long RNA species contain segments of 10-20 nt precisely complementary to conserved regions in rRNA. In the case of rRNA methylation, the so-called box C/D snoRNAs act to guide a methylation reaction, mediated by a nucleolar protein complex, including fibrillarin as the most prominent member. Similarly, the box H/ACA snoRNAs are guiding the pseudouridinylation of numerous uridines in eukaryotic rRNAs (Decatur and Fournier, 2003).

Transfer RNAs contain large proportion, up to 25 %, of the modified nucleotides. Nearly 80 different types of modifications or hypermodifications at more than 60 different positions in tRNAs have been identified. Hypermodified nucleotides, such as i^6A (N^6-isopentenyladenosine), are usually found adjacent to the 3´ nucleotide of the anticodon, when this one is A or U. It is widely accepted that the low polarity of such hypermodified bases strengthens the otherwise relatively weak base-pairing associations of the anticodon´s 3´ base with the codon. Furthermore, some of the tRNA modifications form important recognition elements for aminoacyl-tRNA synthetases, enzymes responsible for attaching the correct amino acid to a tRNA. Even the secondary structure elements in tRNA, like D loop and TΨC loop, are named after modified nucleotides dihydrouridine (D) and pseudouridine (Ψ), respectively (Bjork et al., 1987).

1.3 RNA editing

RNA editing as a term is used to describe all the molecular processes which lead to the change in the information content of an RNA molecule on a single nucleotide level with respect to the information encoded in DNA. The process itself occurs either by enzymatic modification of nitrogenous bases of nucleotides or by the insertion or deletion of short nucleotide stretches. This type of minor posttranscriptional RNA modification has been observed in tRNA, rRNA, and mRNA molecules of eukaryotes, but not prokaryotes.

1.3.1 Insertion and deletion RNA editing

The first reported case of RNA editing was observed in the mitochondria of the parasitic kinetoplastid protozoan *Trypanosoma brucei* (Benne et al., 1986). The mechanism proceeds through binding of 50-70 nt trans-acting guide RNAs (gRNAs) to the newly transcribed mRNA. This class of small RNAs provides sequence-specific information where editing must occur because of the complementarity of its 5´ anchor sequence to its cognate mRNA just upstream of the first editing site. On the other hand, the central portion of trans-acting gRNA is complementary to the mature edited mRNA sequence. These insertions or deletions of the non-genomically encoded uridines in the coding regions of the protozoan mitochondrial mRNAs create methionine translation initiation codons, correct frameshift mutations or even give rise to new open reading frames (Alfonzo et al., 1997; Feagin et al., 1988; Maslov and Simpson, 1992). Remarkably, this mechanism has been found to be absolutely essential for producing functional proteins. Since this first discovery, many mRNA precursors of several kinetoplast protozoa have been found to be edited by insertion or, rarely, deletion of uridine residues (Estevez and Simpson, 1999; Gott and Emeson, 2000).

1.3.2 Ribonucleotide deamination

As previously described, RNA editing includes both the insertion/deletion of nucleotides and the deamination of their nitrogenous bases. The latter type, although more recently discovered, is the more widespread type of RNA editing and includes the well-characterized conversion of cytidine (C) to uridine (U) and adenosine (A) to inosine (I) by hydrolytic deamination of C4 in pyrimidine and C6 in purine rings, respectively, as depicted in Figure 1.2.

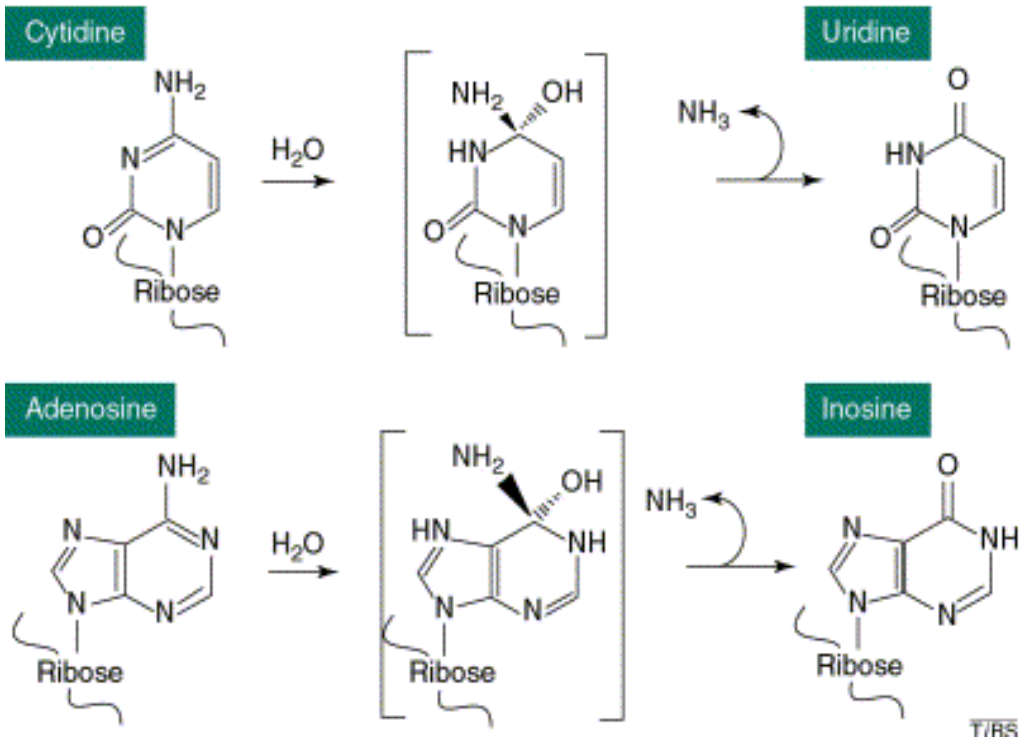

Figure 1.2: Cytidine and adenosine deamination

The reaction mechanism of cytidine and adenosine conversion includes a nucleophilic attack of an activated water molecule at the C4 or C6 of the respective rings in an intermediate step. Subsequently, an ammonia molecule is released, resulting in a stable uridine or inosine, respectively (Gerber and Keller, 2001).

1.3.2.1 Cytidine to uridine (C-to-U) RNA editing

Cytidine deamination in mRNA encoding the mammalian intestinal apolipoprotein B (ApoB) was the first observed RNA editing event of the deamination type in vertebrates (Chen et al., 1987; Powell et al., 1987). APOBEC1, the enzyme responsible for this site-specific alteration belongs to the APOBEC protein family and represents the catalytic subunit of the macromolecular complex known as editosome.

APOBEC1, however, is not capable of carrying out the deamination reaction without additional factors. *In vitro*, a protein complex formed between APOBEC1 and an auxiliary protein factor, APOBEC1 complementation factor (ACF), was shown to be necessary and sufficient for the ApoB mRNA editing (Mehta et al., 2000).

ApoB as a component of plasma lipoprotein particles is essential for the transport of cholesterol and triglycerides in the plasma. The protein exists in two versions in different tissues, each one involved in its distinct lipoprotein metabolism pathway (Chan, 1992). In the liver, the pre-mRNA encodes for the full-length protein, ApoB-100, which is part of the VLDL complex (Very low density lipoprotein) required for the transport of endogenously synthesized cholesterol and triglycerides in the blood stream. During lipid turnover, VLDL particles indirectly give rise to LDL particles (low density lipoprotein) whose increased plasma level is coupled with higher risk for developing atherosclerotic conditions (Herz and Willnow, 1995).

In small intestine, on the other hand, a shorter version of ApoB, ApoB-48, is produced from the same gene locus but with the C-terminal truncation. This protein version contributes to synthesis and secretion of chylomicrons, large lipoprotein complexes of the intestines. These huge particles cannot be converted into LDL particles and, therefore, represent lower risk factor for acquiring cardiovascular diseases. ApoB-48 version of the protein arises due to specific deamination of a cytidine residue at the position 6666. This change gives rise to the premature stop codon (UAA) instead of glutamine codon (CAA) and, thus, results in the synthesis of a shorter protein (Chen et al., 1987; Powell et al., 1987).

The APOBEC family members are also capable of deaminating cytidines in single-stranded DNA, this function being of high importance in the regulation mechanisms of the vertebrate immune system. Firstly, introduction of C-to-U hypermutations in the negative strand of the human immunodeficiency virus (HIV) DNA is mediated by the two APOBEC family members: APOBEC3F and APOBEC3G

(Harris and Liddament, 2004). Lentiviruses, however, have evolved the *Vif* gene in order to counteract this effect: Vif protein interacts with APOBEC3G and triggers the ubiquitination and degradation of APOBEC3G via the proteasomal pathway (Donahue et al., 2008; Sheehy et al., 2002).

Additionally, activation-induced (cytidine) deaminase (AID) is a 24 kDa enzyme that also removes the amino group from cytidines in DNA. This enzyme exhibits *in vitro* activity on single-stranded DNA and has been shown to require active transcription in order to exert its deaminating activity (Bransteitter et al., 2003). AID is currently thought to be the main regulator of secondary antibody diversification. It is involved in the initiation of three separate immunoglobulin (Ig) diversification processes: somatic hypermutation (SHM), class switch recombination (CSR) and gene conversion (GC). During somatic hypermutation, base substitution mutations are introduced into the segments coding for the antigen binding region of vertebrate immunoglobulin genes. Furthermore, the constant regions of the immunoglobulin genes can also be replaced in a process known as class switch recombination. AID is indispensable for both of these processes, but the exact mechanism still remains to be elucidated (Harris and Liddament, 2004).

1.3.2.2 Adenosine to inosine (A-to-I) RNA editing

Similar to cytidine deamination, adenosines can be converted into inosines (Figure 1.2). In higher organisms, this is the most prevalent type of RNA editing since inosines are found in many RNAs, for example in human central nervous system. The enzymes responsible for this posttranscriptional modification belong to the family of adenosine deaminases. Their substrates are mostly double-stranded RNAs with loops and bulges and the mechanism of conversion consists of flipping the target adenosine out of the duplex and subjecting it to hydrolytic deamination. The resulting inosine does not posses adenosine´s base-pairing ability for uridine in the opposing strand, which can lead to destabilisation of the double-stranded structure.

Thus far, editing events have been found in both coding and non-coding regions of protein-coding genes, in viral transcripts, in tRNAs, as well as in miRNAs (Athanasiadis et al., 2004; Cattaneo et al., 1988; Levanon et al., 2004; Li et al., 2009; Luciano et al., 2004; Morse et al., 2002; Polson et al., 1996).

Inosines are often encountered in the first anticodon position (wobble position 34) in a number of tRNAs in higher eukaryotes and yeast, but also in $tRNA^{Arg2}$ of

prokaryotes and chloroplast. This is why it is believed that inosines play an essential role in protein synthesis. Their presence at this critical position gives a certain amount of flexibility to the otherwise rigid process of codon triplet recognition, allowing the formation of alternative base pairs with U, C or A in the third codon position (Sprinzl et al., 1998). In addition, N^1-methylinosine (m^1I_{37}) is found at position 37 of $tRNA^{Ala}$ (A37), adjacent to the anticodon. In this case, adenosine deamination occurs first, followed by the methylation of the inosine (Grosjean et al., 1996).

The search for the enzyme responsible for this tRNA modification in yeast lead to the discovery of an open reading frame highly homologous to the already discovered adenosine deaminases in the 3´ region of their translated sequence. Indeed, the enzyme termed Tad1p (tRNA specific adenosine deaminase) or scADAT1 (*Saccharomyces cerevisiae* adenosine deaminase that acts on tRNA), was demonstrated to specifically target A37 of $tRNA^{Ala}$ (Gerber et al., 1998) and was subsequently cloned from human (Maas et al., 1999), mouse (Maas et al., 2000) and from *Drosophila melanogaster* (Keegan et al., 2000). Some time after, two additional deaminases capable of triggering the conversion of A34 into an inosine have been described in yeast: Tad2p/scADAT2 and Tad3p/scADAT3. The two enzymes form a catalytically active heterodimer. Contrary to ADAT1, ADAT2 and ADAT3 yeast knock out strains are lethal, emphasising the importance of this editing mechanism. In contrast to other adenosine deaminases, ADATs do not contain an RNA binding motif, indicating that the deaminase domain directly recognises its substrate (Gerber and Keller, 1999; Gerber and Keller, 2001; Keegan et al., 2001).

Another, more complex subfamily of adenosine deaminases, adenosine deaminases that act on RNA (ADARs), is thought to have evolved from the ADAT subfamily through the acquisition of double-stranded RNA-binding domains (dsRBDs) and slight alterations of the primary structure in the catalytic (deaminase) domain. Various aspects of ADAR function and activity are dealt with in this work and will be thoroughly explained in the following sections.

1.4 ADARs (adenosine deaminases that act on RNA)

1.4.1 First reports on ADARs

Changes in the structure of double-stranded RNAs formed between endogenous mRNAs and complementary antisense RNAs injected into *Xenopus laevis* oocytes were simultaneously observed by two research groups in 1987. The RNAs in question were also exhibiting altered mobility shifts on native RNA gels and an increased sensitivity to single-stranded RNA-specific RNAses (Bass and Weintraub, 1987; Rebagliati and Melton, 1987). The observed unwinding activity present in oocytes precluded the silencing of target mRNAs. Later on, it was reported that the failure to rehybridise is caused by the irreversible conversion of many adenosines into inosines. As a result, A-U base pairs are converted into I-U mismatches, leading to the destabilisation and the “unwinding” behaviour of the RNA duplex (Bass and Weintraub, 1988; Polson et al., 1991; Wagner et al., 1989).

The enzyme responsible for this covalent modification was first identified in frogs (Hough and Bass, 1994) and shortly after in mammals (Kim et al., 1994a; O'Connell and Keller, 1994).

Since the same RNA modifying activity was identified in two different laboratories, two names for the same enzyme, dsRAD (dsRNA adenosine deaminase) and DRADA (dsRNA adenosine deaminase), were published. Shortly after, a second enzyme with similar capacity to covalently modify adenosines to inosines in the dsRNA was identified, namely RED 1 (dsRNA specific editase 1) (Melcher et al., 1996). In order to avoid the confusion in the nomenclature system for the newly discovered enzyme family, a conclusion was reached to use ADAR1 and ADAR2 as the standardised names for the proteins formerly known as dsRAD/DRADA and RED1, respectively (Bass, 1997).

Since then, ADARs have been found in animals only, but in organisms as evolutionarily divergent as *Caenorhabditis elegans* and *Homo sapiens* and have been shown to be essential for their development.

1.4.2 Members of the ADAR family

As previously mentioned, ADARs are proteins specific for the animal kingdom, with one family member encoded in *Drosophila melanogaster* genome, two in *Caenorhabditis elegans* and three members found in mammals (Figure 1.3).

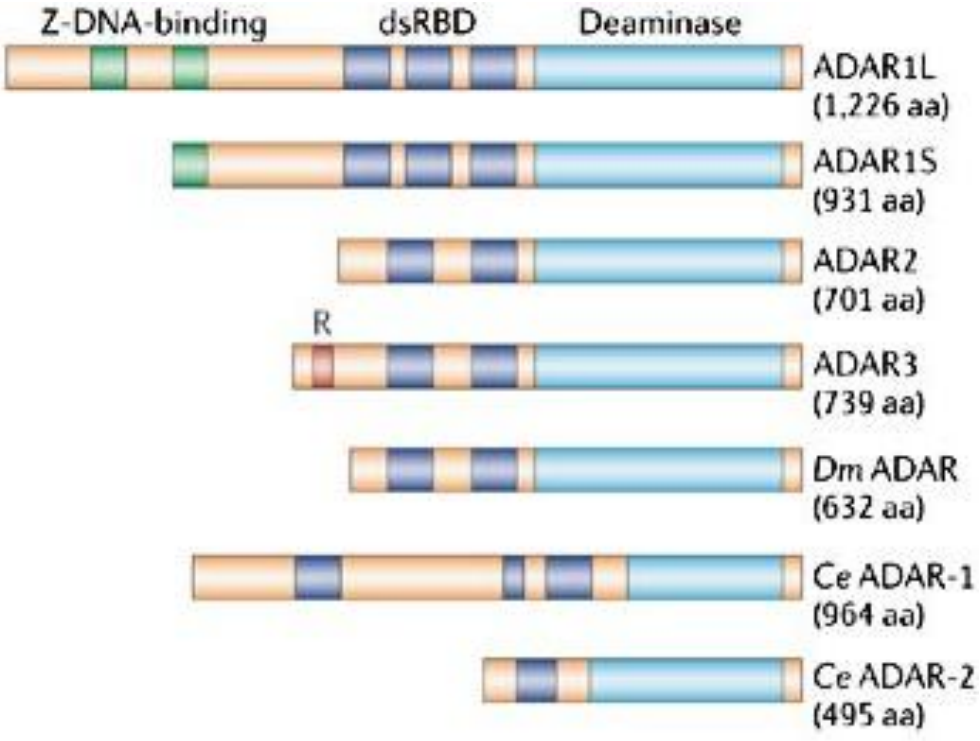

Figure 1.3: ADAR family of enzymes

Three human ADAR family members (ADAR1–3), *D. melanogaster* (Dm) ADAR and two *C. elegans* (Ce) proteins, ADAR-1 and ADAR-2, share common functional domains like dsRBDs and C-terminal deaminase domain (Nishikura, 2006).

A single ADAR gene of the fruitfly is strongly expressed in the central nervous system of embryos and the brains of adult insects (Ma et al., 2001; Palladino et al., 2000b). The protein, termed dADAR, consists of two dsRBDs and C-terminal catalytic deaminase domain. It forms dimers on its substrates and the dimerisation occurs through contacts between N-terminus and the first dsRBD (Gallo et al., 2003). Almost all of its targets localise to the nervous system and comprise voltage- and ligand-gated ion channels, as well as the members of the synaptic release complex (Hoopengardner et al., 2003). Another curious feature of this enzyme is the negative auto-regulation mechanism it employs: its own mRNA serves it as a substrate, where

a specific site in the deaminase domain is edited, leading to the synthesis of the protein isoform with the reduced deamination activity (Keegan et al., 2005; Palladino et al., 2000a).

In *C. elegans*, two ADAR proteins are found. Adr1 has two dsRBDs and a functional catalytic domain, while Adr2 has a single dsRBD, with inactive deaminase domain due to absence of the conserved glutamate (Glu) residue required for its enzymatic activity (Hough et al., 1999; Tonkin et al., 2002). Similar to *D. melanogaster*, identified substrates are important for the functionality of the nervous system (Morse et al., 2002).

In mammals, there are three proteins that belong to the ADAR family: ADAR1, ADAR2 and ADAR3.

Two isoforms of ADAR1 are found in humans, because the ADAR1 gene locus has two distinct promoter sequences, a constitutive and an interferon-inducible one, which leads to the synthesis of two proteins differing in their N-termini. The shorter form, giving rise to a 110 kDa protein and transcribed from the constitutive promoter, is termed ADAR1-c. This isoform is lacking the translation product of the exon 1, because methionine at position 296 is used as a start codon (George and Samuel, 1999; Patterson and Samuel, 1995). In contrast, the longer interferon-inducible variant, named ADAR1-i, starts at methionine 1 and predominantly localises to the cytoplasm where the 150 kDa protein is implicated in viral RNA editing (Wong et al., 2003). Additionally, novel alternatively spliced versions of ADAR1 have been detected (Liu et al., 1997), indicating the importance of ADAR1 regulation on the transcriptional level, as well. ADAR1-i, but not ADAR1-c, has the N-terminal Z-DNA binding domain composed of the Zα and Zβ moieties. This domain is thought to play a role in guiding ADAR1-i to the active transcription sites where, due to unwinding, the DNA adopts a Z-like helical conformation (Herbert and Rich, 2001). The central part of the proteins comprises three dsRBDs, while the catalytic deaminase domain can be found at the C-terminus.

ADAR2, on the other hand, is a much shorter protein of 80 kDa, but it exhibits high level of structural similarity to ADAR1. This comes as no surprise knowing that the protein has 31% sequence identity to ADAR1 with the deaminase domain being the most conserved one (Melcher et al., 1996). Importantly, ADAR2 contains two instead of three dsRBDs and has no Z-DNA binding domains (ZBDs). While ADAR1 is expressed ubiquitously, explaining the observation that inosines can be found in

most tissues, ADAR2 is predominantly expressed in the brain, which correlates with the highest abundance of inosines in the brain (Paul and Bass, 1998).

The most recently discovered member of the human ADAR family, 81 kDa ADAR3 shows an exclusive brain expression pattern, particularly in the amygdala and the thalamus (Mittaz et al., 1997). Its sequence is 50 % identical to ADAR2, with two dsRBDs and a conserved deaminase domain. Interestingly, an additional N-terminal arginine-lysine-rich sequence (R-domain) of ADAR3 has been identified as a single-stranded RNA binding domain (Chen et al., 2000). However, the enzyme is believed to lack any catalytic activity, firstly because no substrate of ADAR3 has been identified so far and, secondly, because it is unable to edit any of the known ADAR1 or ADAR2 substrates. Even though it is not capable of performing the catalysis, ADAR3 could bind potential substrates of ADAR1 and ADAR2 and, thus, might serve as a negative regulator of the two enzymatically active deaminases (Chen et al., 2000).

Additional ADAR-like gene in vertebrates, TENR, is expressed in the male germline and has only one dsRBD (Hough and Bass, 1997). TENR is missing a key glutamate residue in the active site of the deaminase domain and also lacks zinc-chelating residues. The protein itself, however, has not yet been characterized (Keegan et al., 2004).

1.4.3 Domain architecture of ADARs

All proteins belonging to either ADAR or ADAT family contain a conserved catalytic deaminase domain in their C-terminus, but show great variation in their central and aminoterminal regions.

The deaminase domain bears high homology to the catalytic centres of DNA methyltransferases and cytidine deaminases that act on RNA (CDARs) (Hough and Bass, 1997; Kim et al., 1994b). It is composed of characteristic sequence motifs required for the catalysis of hydrolytic deamination. One of them includes highly conserved glutamate residue and is believed to mediate the proton transfer during catalysis. Additionally, conserved histidine and cysteine residues are indispensable for the complexation of a zinc ion, thereby constructing the active site of the enzyme. Similar to most metalloenzymes, a water molecule is activated by the zinc ion, enabling its oxygen atom to nucleophilically attack C6 atom of the adenosine´s purine

ring (Lai et al., 1995). The affected adenosine is flipped out of the RNA helix due to spatial reasons during the editing reaction (Stephens et al., 2000). Interestingly, the dsRBDs of ADARs might facilitate this conformational change through enhancement of the RNA duplex flexibility (Yi-Brunozzi et al., 2001).

In addition to the deaminase domain, only the interferon-inducible isoform of ADAR1 (ADAR1-i) harbours two Z-DNA binding domains (Zα and Zβ) in its N-terminus. Structural analyses have shown that each domain builds up a helix-loop-helix structure similar to the one found in DNA-binding proteins, and can bind left-handed, supercoiled DNA Z-helix *in vitro* (Herbert et al., 1997; Schade et al., 1999; Schwartz et al., 1999). Although this type of DNA double helix has not been identified *in vivo* yet, it is speculated that the progressing RNA polymerases might leave such an unusual structure behind while performing transcription, which led to the proposition that the ZBDs direct ADAR1-i to the transcriptionally active sites (Herbert et al., 1997). For *Xenopus laevis* ADAR1, in contrast, it has been demonstrated that the ZBDs are dispensable for chromosomal localization (Eckmann and Jantsch, 1999). Furthermore, there are hypotheses suggesting the importance of ZBDs for viral dsRNA binding in the cytoplasm, required for hyperediting in the antiviral response (Brown et al., 2000; Herbert et al., 1997).

While ADAR1 contains three double-stranded RNA-binding domains (dsRBDs) in its central region, only two dsRBDs can be found in ADAR 2 and ADAR 3. They are involved in the substrate recognition and binding, but also serve as modulators of the intracellular trafficking of ADAR1 (Fritz et al., 2009; Strehblow et al., 2002).

1.4.4 Substrate recognition by ADARs and the mechanism of adenosine deamination

Although certain sequence contexts serve as preferred sites for ADAR binding and editing, substrate specificity of ADARs is mainly a matter of RNA structure, and not its sequence. Currently, it is poorly understood how the enzyme discriminates between adenosines that have to be converted into inosines and those that are not altered. Albeit, a longer double-stranded RNA helix interrupted by mismatches, bulges and loops represents an ideal ADAR-binding platform (Bass, 2002). However, there is evidence that the substrate recognition and catalysis by ADARs could involve the base that opposes the edited adenosine. In this case, adenosines in A:C

mismatches were edited at much higher rate by both ADAR1 and ADAR2 when compared to A:A or A:G mismatches, or even A:U base pairs at the same location within dsRNA (Wong et al., 2001). Also, the reports on 5´ nearest neighbour preference have been published (Dawson et al., 2004; Riedmann et al., 2008).

The exchange of the deaminase domains between ADAR1 and ADAR2 resulted in the substrate specificity typical for the enzyme whose deaminase domain was part of the protein chimaera (Wong et al., 2001). However, studies on chimaeric proteins containing two dsRBDs of the RNA-dependent protein kinase (PKR) combined with the ADAR1 deaminase domain revealed the importance of ADAR dsRBDs in the substrate recognition, as well (Liu et al., 2000).

To give further insight into how the substrate specificity is achieved, an NMR chemical shift perturbation study was performed using the two dsRBDs of rat ADAR2 and the GluR-B R/G site stem-loop as a substrate. The protein and the RNA surfaces involved in the complex formation were identified, with dsRBD1 found to recognize the conserved terminal pentaloop and dsRBD2 found to interact with two bulged nucleotides adjacent to the R/G editing site. It is, therefore, evident that different dsRBDs from the same protein can act cooperatively in order to achieve a productive recognition of the cognate substrate (Stefl et al., 2006).

In order to clarify the exact reaction mechanism by which the adenosine deamination proceeds, crystal structure of the human ADAR2 catalytic domain was solved. The active site is defined by a zinc ion that coordinates a water molecule responsible for the ammonia displacement during the deamination reaction, with the coordination of the zinc ion appearing essentially identical to the geometry seen at the catalytic centers of cytidine deaminase and TadA, a member of the ADAT2 family. The site of nucleophilic attack during the ADAR reaction (C6 of adenine) is buried in the major groove of the dsRNA. Due to its inaccessibility, ADARs are thought to employ a base-flipping mechanism to bend the target adenosine out of the dsRNA A-helix. Consistently, the catalytic zinc centre is located in a deep pocket surrounded by the positively charged residues, making it an ideal binding surface for the negatively charged dsRNA. Additionally, catalytic activity of ADAR2 depends on the incorporation of inositol hexakisphosphate (IP_6) within the protein core. This polyphosphate is considered crucial for the protein-folding stabilization, but it could also be involved in the modulation of ADAR2 enzymatic activity (Macbeth et al., 2005).

1.4.5 ADAR substrates

ADARs in higher vertebrates have thus far been found to target three main types of substrates. Firstly, in mammals, as well as in flies and worms, ADAR-mediated RNA editing regulates important functional properties of neurotransmitter receptors in the central nervous system (Maas et al., 2006). Furthermore, another important role for ADARs in mammals is thought to be editing and hyperediting of viral dsRNA as a defense mechanism against infections. Finally, in recent years the effects of adenosine deamination in the non-coding substrates, most notably the microRNA precursors, emerged as a significant regulatory mechanism in RNA interference (Ohman, 2007).

1.4.5.1 ADAR editing in the central nervous system

The most prominent and best described targets of ADARs in the central nervous system are the pre-mRNAs encoding subunits of the glutamate receptor, the serotonin 2C receptor in mammals or the potassium channel in squids (Burns et al., 1997; Lehmann and Bass, 2000; Melcher et al., 1996; Patton et al., 1997).

L-glutamate is the major excitatory neurotransmitter in the vertebrate central nervous system. There it serves the glutamate-gated ion channels in mediating rapid excitatory transmission and synaptic plasticity (Greger and Esteban, 2007). Structurally, glutamate-gated ion channels represent heterotetrameric cation channels which can be divided into three main subtypes: AMPA (α-amino-3-hydroxy-5-methyl-4-isoxasolepropionic acid) receptors, NMDA (N-methyl-D-aspartate) receptors and kainate receptors, all concentrated at postsynaptic sites. Being largely Ca^{2+}-impermeable and displaying extraordinarily fast kinetics, AMPA receptors are responsible for the primary depolarization in glutamate-mediated neurotransmission. Their functional properties mostly depend on the subunit composition and the modifications introduced through RNA splicing and RNA editing. They assemble from four subunits, termed GluR1-GluR4 or, alternatively, GluR-A – GluR-D. ADAR-mediated editing occurs at multiple sites in the GluR-B subunit pre-mRNA (Seeburg et al., 1998).

The editing site in exon 11 of the GluR-B pre-mRNA has been shown to be essential. Here, RNA editing leads to conversion of a glutamine (Q) codon (CAG) into an arginine (R) codon (CGG) with 99.9 % efficiency. The converted residue is found

in the second transmembrane domain, known to take part in the receptor´s central channel formation. The amino acid composition of the entire pore region is, of course, crucial for the ion transport efficiency, with the assembly of Q/R site-edited subunits into the receptor significantly lowering the channel´s Ca^{2+} permeability (Sommer et al., 1991). This arginine residue has also been shown to decrease the rate of the receptor assembly and its transport to the synaptic membranes in cultured hippocampal neurons (Greger et al., 2003; Greger et al., 2002). In this case, the editing is performed by ADAR2 and the binding platform for the enzyme is formed between the exon 11 and the editing complementary sequence (ECS) located in the downstream intron (Higuchi et al., 1993; Melcher et al., 1996).

The second recoding editing site in GluR-B pre-mRNA is the R/G site, with adenosine deamination occurring at up to 80 % efficiency in three out of four AMPA receptor subunits (GluR-B, -C and –D). It is located in exon 13, at position -2 with respect to the 5´ splice site of intron 13. In this case, an arginine (R) codon (AGA) is turned into a glycine (G) codon (GGA). This site exhibits varying editing levels during development and its physiological role is to increase recovery rates from desensitization (Lomeli et al., 1994).

Editing at both Q/R and R/G site is having a crucial influence on splicing of GluR-B pre-mRNA. Reduced editing levels at Q/R site cause aberrant retention of the downstream intron 11 (Higuchi et al., 2000; Schoft et al., 2007). On the other hand, R/G site is situated only one nucleotide upstream of the intron 13 5´ splice site. Editing at this position leads to a retardation in the removal of the adjacent intron, thereby stimulating the correct alternative splicing decisions in the downstream regions of the transcript (Schoft et al., 2007).

Serotonin (5-hydroxytryptamine) is a neurotransmitter important for neuronal functions such as sleep regulation, appetite or pain. The molecule interacts with a large family of receptors and triggers signaling that is crucial for proper neurotransmission. Indeed, binding of serotonin to its transmembrane receptor is involved in several physiological and behavioral processes such as production of cerebrospinal fluid and regulation of feeding behavior (Maas and Rich, 2000). The human 5-hydroxytryptamine receptor of subtype 2C (5-HT_{2C}R) is ubiquitously distributed throughout the central nervous system and is believed to regulate mood, appetite and sexual behavior (Molineaux et al., 1989). Functionally, it belongs to the class of G-protein-coupled receptors which are linked to the activation of

phospholipase C, an enzyme that triggers synthesis of secondary messengers (Hoyer et al., 1994).

Editing in this primary transcript takes place at five different positions in exon 5, namely editing sites A,B,C,D and E. All modified adenosines are situated within the pre-mRNA segment encoding the second intracellular loop of 5-HT_{2C}R, a receptor domain important for G-protein coupling (Niswender et al., 1998). While editing at sites A and B is accomplished by ADAR1, sites C,D and E are deaminated by ADAR2 (Wang et al., 2000).

Interestingly, in rat at least seven and in humans twelve different isoforms of the 2C subtype of serotonin receptor (5-HT_{2C}R) have been found resulting from alternative editing at five different sites by ADARs. The receptor isoform whose pre-mRNA is edited at all five sites carries amino acid substitutions at positions 156, 158 and 160, where the genomic DNA-encoded isoleucine, asparagine and isoleucine (I,N and I) are converted into valine, glycine and valine (V,G and V), respectively. Strikingly, the resulting amino acid substitutions result in decreased G-protein coupling and less efficient triggering of the phospholipase C signaling cascade. The ligand affinity is also affected, with the unedited version exhibiting the highest and the fully edited version the lowest affinity for the neurotransmitter, while partially edited isoforms show intermediate affinity (Burns et al., 1997; Flomen et al., 2004; Niswender et al., 1998). Mutant mice with the knocked-in pre-edited (VGV) receptor version were showing elevated energy expenditure and loss of fat mass due to constitutive activation of the sympathetic nervous system (Kawahara et al., 2008).

Recently, a third notable target of ADAR editing in the central nervous system was reported in mouse, namely the gamma-aminobutyric acid (GABA) type A ($GABA_A$) receptor subunit α3 (Gabra-3) pre-mRNA. This receptor is the member of the Cys-loop ligand-gated ion channel superfamily and conducts chloride (Cl^-) ions upon GABA-mediated activation. The modified adenosine is found in exon 9 within a short stem-loop structure, opposing the cytidine residue, and leads to I/M (isoleucine to methionine) amino acid substitution in the transmembrane (TM) domain 3 in the subunit α3 of the five-subunit $GABA_A$ receptor. Both ADAR1 and ADAR2 can mediate this developmentally regulated adenosine deamination: 50 % of transcripts are edited in newborn mice at day 2, while those in adult mouse brain are edited to almost 100 % (Ohlson et al., 2007).

1.4.5.2 Editing of viral dsRNA

The full-length isoform of ADAR1, ADAR1-i, is transcribed from a promoter region containing an interferon-stimulated response element (ISRE) (George and Samuel, 1999). Like most genes whose activation is triggered through interferon signaling, it has been implicated in the cellular antiviral defense mechanism. Interferon synthesis is induced by dsRNA, a byproduct of the replication cycles of many RNA viruses and serves as the first line in combating viral infections in mammals (Sen, 2001).

The alterations in cDNA sequences consistent with the ADAR enzymatic activity have been reported for many viruses. The editing of viral RNAs occurs promiscuously in almost every case examined, with up to 50 % of all adenosines modified.

Such hypermutation is seen in viral RNA isolates of host cells invaded by parainfluenza virus, respiratory syncytial virus and vesicular stomatitis virus (Bass, 2002). Pre-mRNA encoding the matrix protein of the measles virus, crucial in viral budding, is also found nonselectively deaminated at adenosine residues. This prevents the progression of the lytic cycle and forces the virus to remain in the persistent infection state (Sheppard et al., 1985). All of the above mentioned pathogens belong to the group of paramyxoviruses whose non-segmented RNA genome requires an RNA-dependent RNA polymerase for the transcription into an mRNA. Another similarity among them is the exclusively cytoplasmic replication cycle, leading the scientists to believe that there is a cytoplasmic endonuclease targeting only hyperedited dsRNAs. Indeed, such an enzyme, Tudor staphylococcal nuclease (Tudor-SN), has been discovered and the experiments have revealed its specificity for the highly edited dsRNAs comprising alternating I-U and U-I weak basepairing partners, leaving the moderately deaminated dsRNAs intact (Scadden, 2005; Scadden and Smith, 2001b).

A limited number of viruses, however, have already evolved different mechanisms to take advantage of the ADAR-mediated RNA editing. During the late stage infection with the polyoma virus, for example, newly synthesized transcripts form double-stranded regions with the complementary early mRNAs, leading to the hyperediting of the latter. While the hyperedited early transcripts become export-incompetent, the late ones reach the cytoplasm and give rise to proteins fundamental

for the transition from early to late stage infection (Bass, 2002). In the nucleus, such hypermutated RNAs, like the polyoma virus early transcripts, are anchored to the nuclear matrix by the inosine-rich RNA-binding protein $p54^{nrb}$, which forms a complex with the splicing factor PSF and the inner nucler matrix protein matrin 3, thereby preventing the nuclear exit of the promiscuously edited messages and, subsequently, their translation to the non-functional proteins (Zhang and Carmichael, 2001).

There are also selective editing events described for viral RNAs, like in hepatitis delta virus (HDV), a subviral human pathogen that requires hepatitis B virus (HBV) for packaging. Its genome encodes a single open reading frame, giving rise to hepatitis delta antigen (HDAg). Due to ADAR-mediated RNA editing, this protein occurs in two forms. The shorter form is expressed constitutively, while the adenosine deamination within the amber (UAG) stop codon gives rise to tryptophane (UGG) codon in so-called amber/W site, resulting in the extension of the open reading frame by additional 19 amino acids (Polson et al., 1996). The shorter protein (HDAg-S) plays a role in the viral genome replication, while the longer isoform (HDAg-L) leads to the replication inhibition and promotes the assembly of new viral particles (Ryu et al., 1992).

Recently, ADAR1 was shown to stimulate human immunodeficiency virus type 1 (HIV-1) replication by using both editing-dependent and editing-independent mechanisms. Over-expression of ADAR1 in HIV-1 producer cells increases viral protein accumulation in an editing-independent manner. Additionally, ADAR1 also associates with HIV-1 RNAs and edits adenosines in the 5' untranslated region (UTR) and the Rev and Tat coding sequences. It is believed that, just like polyoma virus and HDV, HIV-1 has adapted in a way to make use of specific RNA editing activity of the host cell (Doria et al., 2009).

1.4.5.3 Non-coding ADAR substrates

Recent developments in the field of RNA editing indicate that mRNA recoding at specific sites is not the principal function of RNA editing. Extensive bioinformatics screening for novel ADAR substrates revealed a completely new set of putative ADAR targets in Alu elements and among microRNA precursors (Athanasiadis et al., 2004; Blow et al., 2006; Levanon et al., 2004; Li et al., 2009).

Alu elements are found at approximately 1.4 million copies in the human genome and comprise just over 10 % of the entire genome content. These 300 bp

retrotransposon sequences contain inverted repeats that can fold into double-stranded RNAs upon transcription, making them an ideal ADAR target (Athanasiadis et al., 2004; Levanon et al., 2004; Morse et al., 2002). When a closer look was taken on the nature of the newly emerged ADAR substrates, it was determined that 90 % of them resides in one of the members of Alu family. These editing events can cause creation or elimination of splicing signatures and affect the alternative splicing of Alu-derived exons (Athanasiadis et al., 2004; Lev-Maor et al., 2007; Levanon et al., 2004), but can also serve as a global regulatory mechanism in controlling RNA stability, splicing or translation (Nishikura, 2004).

Another new class of non-coding RNAs crucial in the regulation of gene expression was shown to be targeted by adenosine deamination, namely the microRNA precursors. In the first report on the topic it has been demonstrated that the precursor of microRNA-22 (miRNA-22) forms an 85 nucleotide stem-loop that can be bound and edited at up to six sites by ADAR1 and ADAR2 in mammals (Luciano et al., 2004). At the time, it was poorly understood what the purpose of miRNA editing was, but over the years it became clear that such posttranscriptional modification has a profound influence on fate of both microRNAs and their targets. Second reported editing target among microRNA precursors was miR-142 that was exhibiting loss of function after RNA editing, because the cropping of inosine-containing primary transcripts into pre-miRNA-142 by the Drosha-DGCR8 complex was inhibited. Moreover, highly edited primary miR-142 transcripts, unlike unedited ones, were cleaved *in vitro* by Tudor-SN, a cytoplasmic endonuclease, specific for dsRNAs with high inosine content and a component of the RISC complex (Yang et al., 2006). Furthermore, in the study on miR-151 editing, it was reported that the inosine content was increasingly higher in pre-miRNA-151 and mature miRNA-151 in several mouse tissues (Kawahara et al., 2007a). Here, Dicer-TRBP processing of pre-miR-151 is affected when structure-destabilising inosines are present. Finally, miRNA editing can lead to the redirection of mature miRNA to a different mRNA target, like in the case of mature miR-376 where the editing site is found within the seed region crucial for initial interaction steps with the complementary sequence in cognate mRNA (Kawahara et al., 2007b).

As soon as the first reports on miRNA precursor editing were published, the prospect of ADARs accompanying miRNA precursors through nuclear pores out to the cytoplasm began to emerge, mainly because it was not clarified how the edited

primary miRNA transcripts, thought to be export-incompetent, reach the cytoplasm for degradation. ADAR1p150, for example, is an interferon-inducible shuttling protein known for its role in antiviral response through cytoplasmic RNA editing of various pathogens. This full-length isoform of ADAR1, however, is expressed at steady-state levels even in cells and tissues that have never been infected (Patterson and Samuel, 1995). Therefore, it is plausible that this fraction of ADAR1p150 could accompany editing targets among miRNA precursors to the cytoplasm. Furthermore, being a dsRBD protein, ADAR1p150 could exploit miRNA precursor as an RNA bridge in the formation of export-competent ternary complex with Exportin-5. Exportin-5 is the karyopherin identified as an export receptor for pre-miRNAs, as well as for other small structured RNAs like viral minihelix RNAs or tRNAs (Calado et al., 2002; Gwizdek et al., 2003; Lund et al., 2004; Yi et al., 2003). This importin-β family member has first been identified as a nuclear export factor for ILF3, dsRBD-containing protein (Brownawell and Macara, 2002). Later on, it was shown to mediate RNA-dependent export of a number of dsRBD proteins, like JAZ, Staufen-2 or eEF1A (Calado et al., 2002; Chen et al., 2004; Macchi et al., 2004). Remarkably, dsRBDs of cargo proteins are in most cases indispensable for the interaction with Exportin-5. In addition, this interaction is thought to be mediated, or at least strengthened, through an RNA bridge (Gwizdek et al., 2004).

Exportin-5 was also shown to interact with dsRBDs of ADAR1 and interfere with its nuclear import via Transportin-1, implying its role in ADAR1 nucleocytoplasmic distribution (Fritz et al., 2009; Poulsen et al., 2001; Strehblow et al., 2002). Together with the role of Exportin-5 in pre-miRNA nuclear export and ADAR1-mediated editing of miRNA precursors, this prompted us to address the question of ADAR1 association with miRNA precursors in both nucleus and cytoplasm and, possibly, their nuclear co-export in the second part of this thesis.

1.4.6 Biological importance of ADARs

ADARs have been identified in all metazoan species, including *C. elegans, D. melanogaster, X. laevis* and humans (Bass, 1997). The biological relevance of the enzymes was investigated by generating knock-out animals in several species. Although the invertebrate species remain viable upon ADAR deletion, they exhibit severe phenotypes. ADAR elimination in mammals, on the other hand, results in lethality, proving the adenosine deamination to be essential for normal mammalian development. Altogether, it can be concluded that during the course of evolution the role of ADAR-mediated RNA editing became increasingly important, providing higher vertebrates with yet another sophisticated mode of posttranscriptional gene regulation.

Despite their viability, both *adr-1* and *adr-2* knock-outs in *C. elegans* have difficulties detecting and tracking chemoattractants, a phenotype described as aberrant chemotaxis (Tonkin et al., 2002). Strikingly, this phenotype could be partially rescued when animals deficient in one of the *adr* genes were crossed with the strains mutated in one of the genes responsible for the effective RNAi response. These findings, together with reports on low Dicer cleavage efficiency of hyperedited dsRNAs *in vitro*, have provoked a setting of a theory according to which RNA editing and RNAi act as two antagonistic processes (Scadden and Smith, 2001a; Tonkin and Bass, 2003).

A-to-I RNA editing has also been reported for several *D. melanogaster* ion channel gene transcripts. The deletion mutants of the responsible enzyme, dADAR, that lack deamination activity in extracts lack all known site-specific RNA editing (25 sites in three ion channel transcripts). Although *D. melanogaster* larvae homozygous for the deletion in their *adar* locus have no phenotypical changes, adults lacking dADAR show no morphological abnormalities, but exhibit extreme behavioural deficits including temperature-sensitive paralysis, locomotor uncoordination and tremors and, finally, changes in grooming and mating behaviour. Neurodegeneration, affecting particularly the retina and causing ubiquitous brain lesions, accompanies the increase in phenotypic severity. This study provided clear evidence that pre-mRNA editing in *D. melanogaster* alters the nervous system targets to affect adult nervous system function, integrity, and behaviour (Palladino et al., 2000b).

In contrast to severe but viable phenotypes of invertebrate ADAR knock-outs, such deletions in mammals have been shown to be lethal. In the case of ADAR1 locus deletion, heterozigosity lead to birth of apparently completely healthy mice. Homozygous deletants, on the other hand, die between days 11.5 and 12.5. Death occurs because of the halted developmental processes, distorted fetal liver structure and widespread programmed cell death in tissues including heart and vertebra. The apoptotic events, however, are most severely expressed in both hepatocyte and haematopoietic cell lineages of the fetal liver (Hartner et al., 2004; Wang et al., 2004a; Wang et al., 2000).

Recently, experiments with the inducible gene disruption in mice have shown that ADAR1 is not only essential for the maintenance of both fetal liver and adult bone marrow haematopoietic stem cells, but it can also act as a crucial suppressor of interferon signaling and, thus, protect cells from the deleterious consequences of the activation of interferon-stimulated pathways. The latter role of ADAR1 is attributed to the editing of as yet unidentified transcripts involved in the interferon-induced cellular response (Hartner et al., 2009).

Point mutations in locus coding for ADAR1 are associated with dyschromatosis symmetrica hereditaria (DSH), a disease that leads to altered pigmentation pattern of the skin on hands and feet. Patients display only a mild heterozygous dominant phenotype, with their mental and physical condition seemingly unaffected by the mutation (Miyamura et al., 2003).

Similar to ADAR1 mutants, mice heterozygous for the ADAR2 deletion/disruption are viable and appear normal. Their homozygous mutant counterparts, however, die not long after birth, between postnatal days 0 and 20, exhibiting severe neurological symptoms, among them frequent epileptic seizures. This phenotype is a consequence of underediting at exclusively ADAR2-targeted Q/R site of the GluR-B pre-mRNA, giving rise to incorrectly spliced mRNA and, subsequently, receptors with increased Ca^{2+} permeability in hippocampal neurons, leading to hippocampal neurodegeneration. Interestingly, the viability and completely normal phenotype are reverted once both alleles for this underedited transcript get substituted with the genomically pre-edited ones, i.e. where target adenosine is mutated to guanosine. This observation defines the glutamate receptor subunit B pre-mRNA as the most important, if not the only essential, ADAR2 target *in vivo* (Higuchi et al., 2000).

ADAR2 dysfunction in humans has been implicated in couple of serious conditions, most of them unsurprisingly accompanied by neurological and psychiatric symptoms, considering that ADAR2 expression is highest in the brain where most of its substrates reside. Amyotropic lateral sclerosis (ALS), for example, is an illness characterized by the degeneration of motor neurons, leading to the muscular atrophy and respiratory failure. Impaired ADAR2 activity has been implicated in ALS because the editing at Q/R site of GluR-B pre-mRNA is incomplete in 56 % of ALS patients and the corresponding symptoms can be attributed to glutamate excitotoxicity and the excessive Ca^{2+} influx (Kawahara et al., 2004; Kwak and Kawahara, 2005).

5-HT_{2C} receptor pre-mRNA, on the other hand, is edited at five positions and, considering the role that serotonin plays in the central nervous system, it is believed that reduced or disrupted ADAR activity could play a role in psychiatric disorders like depression or schizophrenia. In comparison with healthy individuals, the 5-HT_{2C} receptor pre-mRNA editing pattern is significantly changed in suicide victims with documented chronic major depression. These patients had overedited E and C sites and underedited D site when compared to samples of individuals without similar condition. Since all of the affected sites are ADAR2 targets, altered function of this protein seems to be a likely cause of the symptoms. Moreover, studies on mice have shown that the treatment with the antidepressant fluoxetine leads to exactly reverse changes in editing frequency at E, C and D sites when compared to the suicide victims, implying one of the possible mechanisms of action for this type of medication (Gurevich et al., 2002).

1.5 RNA splicing

One of the main features of the protein-coding primary transcripts in higher eukaryotes is the existence of multiple non-coding intervening sequences known as introns. These long stretches of nucleotides interrupt the expressed sequences, named exons, and need to be precisely removed in order to obtain fully mature, export- and translation-competent message in a process called RNA splicing. All of the reactions leading to intron removal and exon fusion take place in a spliceosome, a macromolecular assembly characterized by a highly dynamic network of RNA-protein and RNA-RNA interactions. The essential components of the spliceosome include five small RNAs – U1, U2, U4, U5 and U6 small nuclear RNAs (snRNAs), all of them in RNA-protein complexes known as snRNPs, small nuclear ribonucleoprotein particle. In addition, there are a plethora of extraspliceosomal protein factors required to make the correct splicing decisions (Guthrie and Patterson, 1988; Maniatis and Reed, 1987).

1.5.1 Sequence requirements in RNA splicing

The exact sites where the enzymatic reactions in RNA splicing take place are defined by the primarily intronic consensus sequences, close to both 5´ and 3´ splice sites. The 5´ end of the intron, the 5´ splice site, is relatively short and contains highly conserved dinucleotide GU at positions 1 and 2 of the intron. In contrast, 3´ splice sites are defined by three separate sequence elements – the branchpoint site, the polypyrimidine tract and the 3´ splice site consensus itself. Together, they make up a loosely defined 3´ splice site region extending up to 100 nucleotides upstream of the actual excision site: the branchpoint adenosine is found surrounded by the short consensus sequence, followed by the polypyrimidine tract and, finally, ending in the 3´ terminal dinucleotide AG of the intron (Padgett et al., 1986).

1.5.2 Chemistry of RNA splicing

All introns within nuclear mRNA precursors are thought to be removed by the same two-step mechanism consisting of two consecutive transesterification reactions (Figure 1.4). The first step involves a direct nucleophilic attack of the 2´-OH of the branchpoint site adenosine on the phosphate at the 5´ splice site, which results in the

displacement of the 3´ oxygen of the 5´ exon and the formation of the 2´-5´ phosphodiester bond between the branchpoint and the 5´-terminal nucleotide of the intron. Therefore, two products of the first transesterification are a detached 5´ exon and an intron-3´exon fragment in a lariat configuration. In the second step, the same 3´-OH of the 5´ exon that had been the leaving group in the first reaction reverts its role and now behaves as a nucleophilic group at the 3´ splice site to similarly displace the 3´ end of the intron with the 5´ end of the 3´ exon. This results in the ligation of 5´ exon with the 3´ exon via new phosphodiester bond and leaves the released intron still in the lariat structure (Padgett et al., 1986; Sharp, 1987).

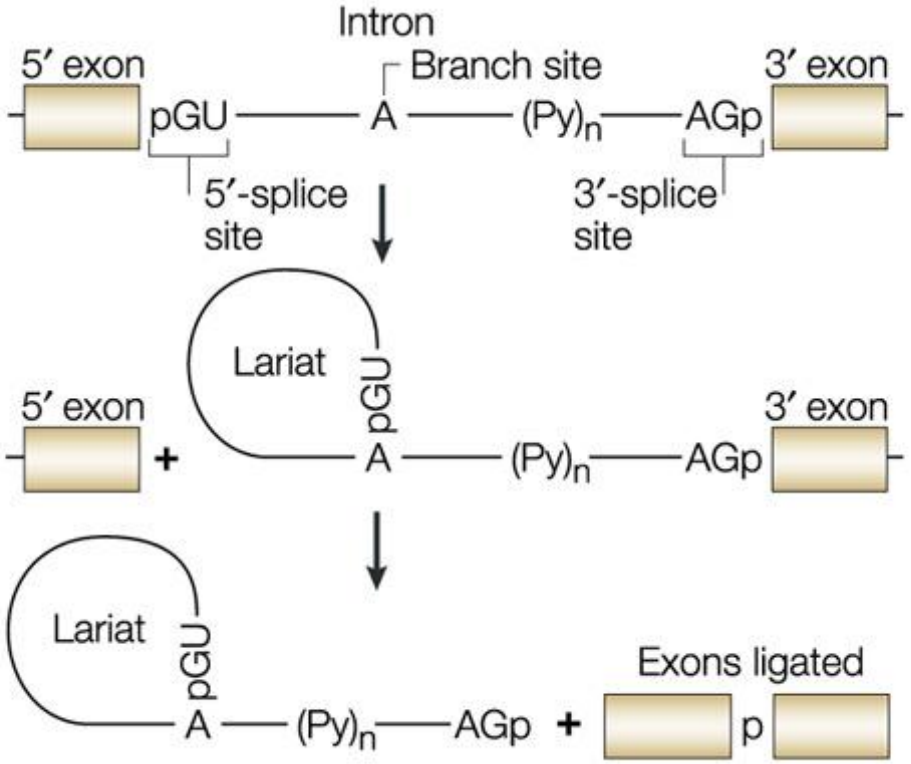

Figure 1.4: Two-step mechanism of RNA splicing

Splicing of pre-mRNA takes places via two transesterification reactions. In the first step, the 2'-hydroxyl group of the branchpoint site adenosine attacks the phosphate at the GU of the 5'-splice site. This leads to cleavage of the 5' exon from the intron and the formation of a lariat intermediate. In the following step, a second transesterification reaction, which involves the phosphate at the 3' end of the intron and the 3'-hydroxyl group of the detached exon, ligates the two exons. This reaction releases the intron, still in the form of a lariat. (Pagani and Baralle, 2004).

1.5.3 Spliceosome assembly

Before two crucial enzymatic reactions of RNA splicing can take place, target pre-mRNA needs to be forced into the splicing-competent configuration, the spliceosome. Immediately after transcription, nascent pre-mRNAs are bound by a multitude of heterogenous nuclear ribonucleoprotein particles (hnRNPs) (Pinol-Roma and Dreyfuss, 1992). However, the earliest event that specifically targets pre-mRNAs to the splicing pathway is the assembly of the early or commitment spliceosomal complex (E complex), where all consensus sequence elements need to be recognized (Figure 1.5). This is achieved through four distinct macromolecular interactions. At the 5´ end of the intron, U1 snRNP stably associates with the 5´ splice site by means of the base-pairing interaction between U1 snRNA and the consensus sequence at the 5´ exon/intron border (Rosbash and Seraphin, 1991). Further downstream, the branchpoint consensus is bound by the splicing factor 1 (SF1), whereas the polypyrimidine tract and the 3´ splice site interact with the 65 kDa and 35 kDa subunits of the heterodimeric U2 snRNA auxiliary factor (U2AF), respectively (Berglund et al., 1997; Wu et al., 1999; Zamore et al., 1992).

Subsequent to the formation of a commitment complex, U2 snRNP joins the growing macromolecular assembly by base-pairing of its U2 snRNA with the branchpoint consensus sequence, marking the formation of the A complex (Parker et al., 1987). Once the U4/U5/U6 tri-snRNP enters the spliceosome, the B complex is established (Konarska and Sharp, 1987). Upon profound rearrangements, including the displacement of the U1 snRNP from the 5´ splice site by the U6 snRNP and the loss of both U1 and U4 snRNPs from the spliceosome, a mature spliceosomal complex is assembled, namely the C complex (Figure 1.5), which catalyzes both reactions required for effective RNA splicing (Bessonov et al., 2008; Makarov et al., 2002).

Once the second step of the splicing is terminated, another macromolecular assembly, the exon-exon junction complex (EJC) accumulates 20-24 nucleotides upstream of exon borders in mature mRNA. The complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay, thereby providing an RNA quality control mechanism to determines whether the RNA species is translation-competent or it should be targeted for degradation (Le Hir et al., 2001).

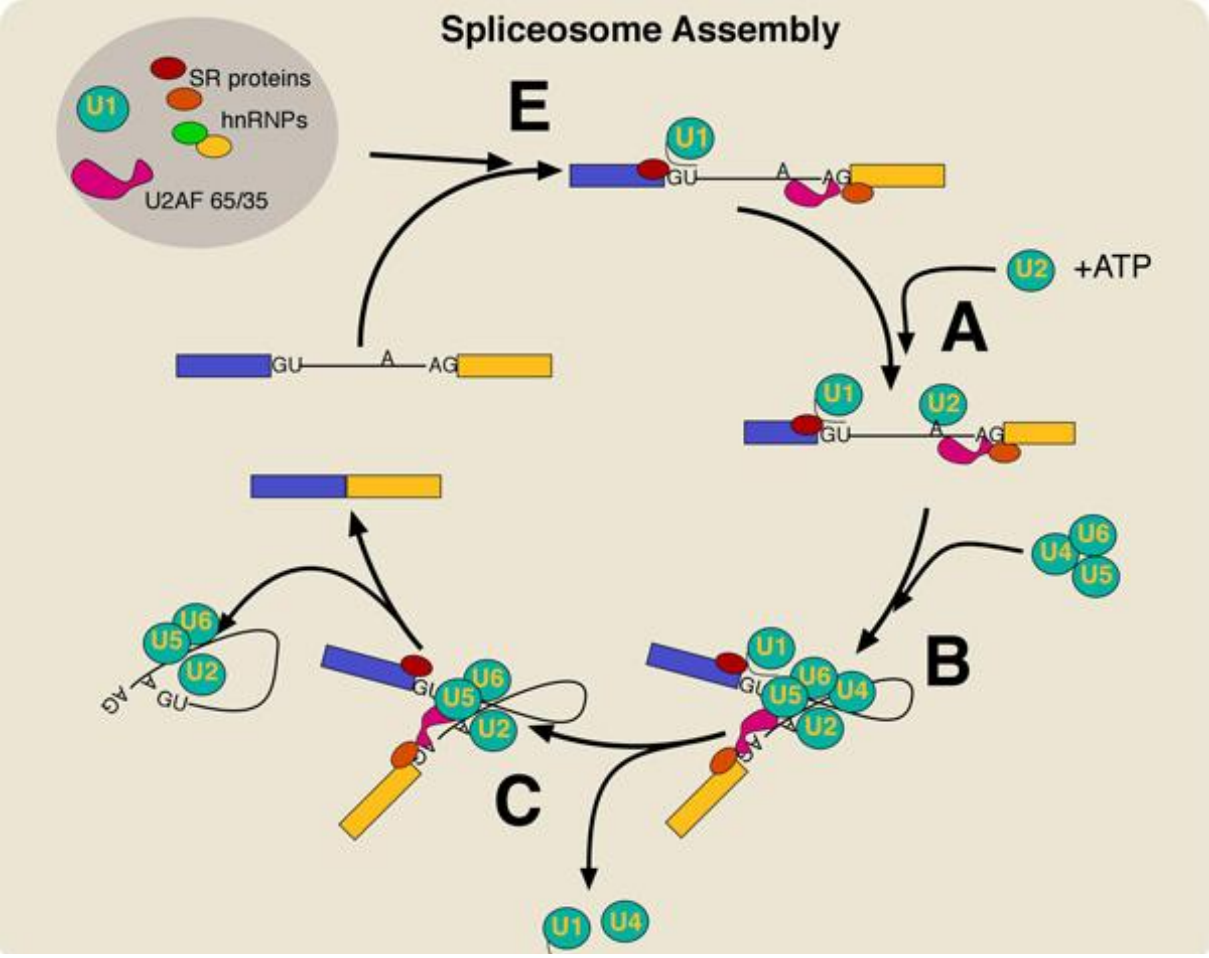

Figure 1.5: Distinct steps of the spliceosome assembly

Complex rearrangements in the structure of the spliceosome take place throughout the process of RNA splicing. E or the commitment complex consists of the U1 snRNP, SF1 and heterodimeric U2AF bound to the intronic consensus sites. Once the U2 snRNP joins, the A complex is formed. Transient B complex is defined through the association of the U4/U5/U6 tri-snRNP with the A complex. The C or the catalytic complex is established after the rearrangements occur in the B complex.

1.5.4 Regulation of RNA splicing

The mechanism of RNA splicing, during which the synthesis of both constitutively and alternatively spliced mRNAs occurs, is regulated by a system of trans-acting factors that bind to cis-acting sites on the pre-mRNA itself. Such trans-acting factors include proteins known as splicing activators that promote the usage of a particular splice site, and splicing repressors that reduce the usage of a particular site. The 5′ splice site, the 3′ splice site and the branch point sequence as three core splicing signals are present in every intron and are also necessary for the splicing reaction, but the exact definition of the exon-intron boundary requires interplay of the cis-acting sequences in pre-mRNAs, known as splicing enhancers and splicing silencers that can map to both exonic and intronic regions (Matlin et al., 2005; Smith and Valcarcel, 2000).

The protein regulators of the RNA splicing can be roughly divided into two groups. First group consists of SR proteins which mostly promote exon inclusion when bound to the cis-acting sequences, while the second group comprises hnRNP family of proteins whose interaction with the cis-acting signals often results in exon skipping (David and Manley, 2008).

SR proteins preferentially interact with both exonic and intronic splicing enhancers (ESEs and ISEs), thereby assisting in the constitutive removal of the adjacent introns. They all have a modular organization and contain at least one N-terminal RNA-binding domain that interacts with the pre-mRNA and a C-terminal RS domain that functions as a protein interaction domain (Graveley, 2000).

Exonic and intronic splicing silencers (ESSs and ISSs), on the other hand, are in most cases bound by one of the members of hnRNP family of splicing factors. This interaction leads to the inhibition of the adjacent splice sites and, consequently, to intron retention (Wang et al., 2004b). The splicing repression at several alternative splice sites occurs through the negative regulation mediated by hnRNP A1 which exhibits strong affinity for the sequences UAGGG$^{A}/_{U}$ and GGCAGGGUGG (Burd and Dreyfuss, 1994; Del Gatto-Konczak et al., 1999; Fairbrother and Chasin, 2000). Some hnRNPs, like hnRNP H, play roles as both splicing repressors and activators, determined by the cis-acting element they bind to (Caputi and Zahler, 2001). Furthermore, the crucial role of hnRNP H in the regulation of alternative splicing of the rat β-tropomyosin gene transcript has been reported (Chen et al., 1999).

In Figure 1.6, the summary of major alternative splicing pathways is depicted (panel A), also indicating the principal interactions that occur between cis-acting elements and trans-acting factors (panel B).

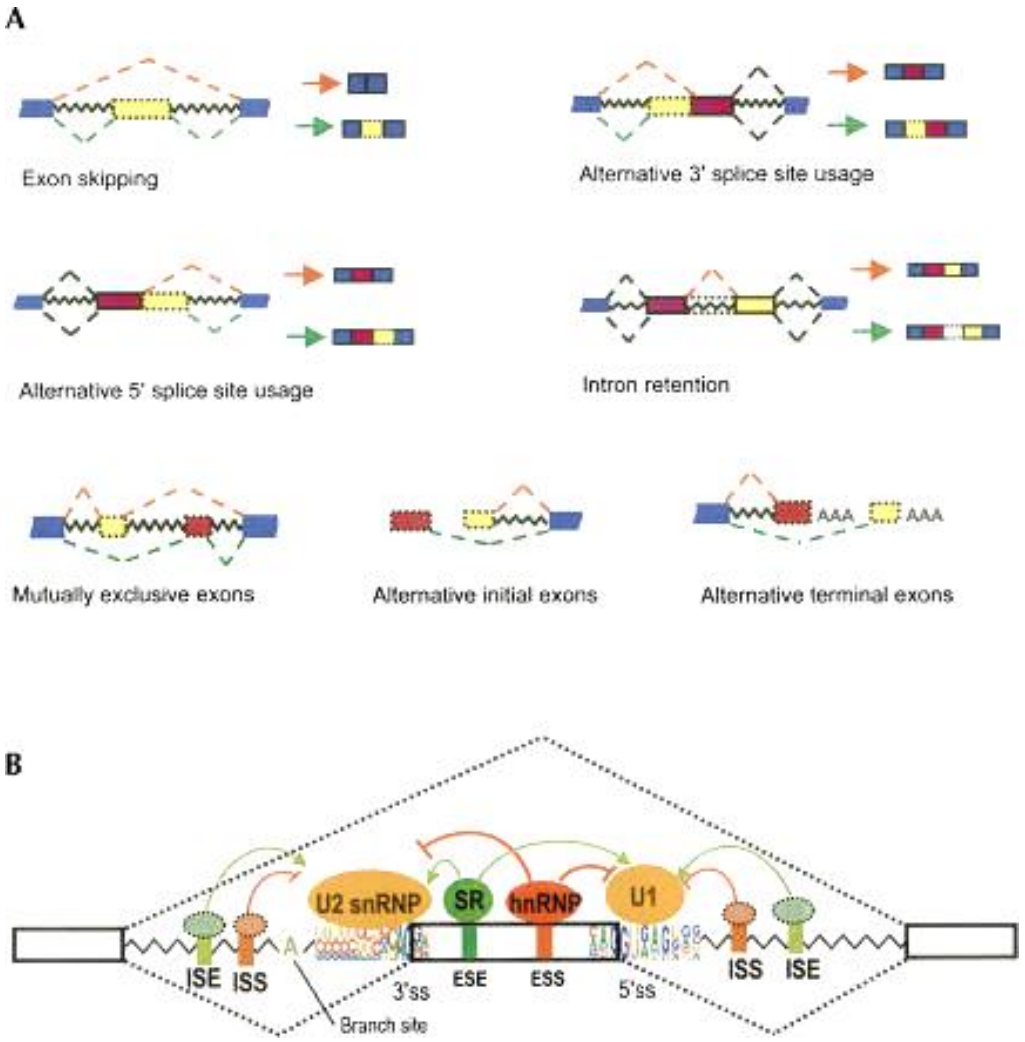

Figure 1.6: Alternative splicing of pre-mRNAs

A: Major alternative splicing pathways. The combination of two or more pathways give rise to more complex alternative splicing network.

B: Regulation of splicing. Exons, introns and splice sites are represented as open boxes, jagged lines and brackets, respectively. The 5´ and 3´ splice site consensus sequences are shown in pictogram with the branchpoint site adenosine indicated. Two alternative splicing pathways (dashed lines) are indicated with the central exon either included or excluded. Regulatory *cis*-elements (ESE, ESS, ISS, and ISE) are spread all along the pre-mRNA and bound by *trans*-acting splicing factors (SR proteins and hnRNPs) (Wang and Burge, 2008).

1.5.5 Connection between RNA editing and splicing

In most cases of ADAR-mediated RNA editing in protein-coding targets, efficient modification of the exonic editing sites relies upon the formation of the suitable double-stranded structure by means of base-pairing with the editing complementary sequence located in the downstream intron (Higuchi et al., 1993). Here, editing needs to occur co-transcriptionally, before the intron is removed through RNA splicing, suggesting that editing and splicing need to be coordinated (Keegan et al., 2001).

There are many reports where RNA editing was shown to be important in making the correct downstream splicing desicions. In *D. melanogaster*, for example, if editing occurs upstream of alternative splice site, a strong correlation between editing efficiency and splice site selection is observed. In contrast, when editing occurs downstream of alternative splicing, no correlation is seen (Agrawal and Stormo, 2005).

Auto-editing of ADAR2 pre-mRNA is another phenomenon where the importance of intimate interplay of editing and splicing is strongly emphasised. In rats, sequence analysis of ADAR2 genomic DNA revealed the presence of adenosine-adenosine (AA) and adenosine-guanosine (AG) dinucleotides at the proximal and distal alternative 3' splice sites of intron 4, respectively. The proximal 3' splice site recruitment depends upon the ability of ADAR2 to edit its own pre-mRNA, converting the intronic AA to an adenosine-inosine (AI) dinucleotide which then mimics the highly conserved AG dinucleotide usually found at 3' ends of introns. Introduction and use of this alternative splice site leads to premature translation termination and the synthesis of non-functional enzyme (Rueter et al., 1999).

Furthermore, RNA editing of an intronic adenosine is crucial for the correct splicing of the pre-mRNA of the serotonin receptor 2C subunit because it results in the structure destabilisation around the adjacent 5´ splice site, facilitating the access of the spliceosomal components (Flomen et al., 2004). Previously, alternative splice variants of this transcript caused by underediting were found enriched in malignant gliomas, most probably because of the synthesis of the truncated non-functional version of the receptor subunit (Maas et al., 2001).

In GluR-B pre-mRNA, RNA editing is crucial for the correct processing, most importantly the removal of introns 11 and 13, with the latter being coupled to proper

downstream alternative splicing decisions (Schoft et al., 2007). At the Q/R site in exon 11, ADAR2-mediated editing is coupled to the editing of the intronic editing hotspots in intron 11, at positions +60 and +262/263/264, respectively. This deamination event is indispensable for the proper removal of intron 11, because ADAR2 knock-out mice show an accumulation of intron 11-containing RNAs in the nucleus, although every other intron is regularly eliminated (Higuchi et al., 2000).

Splicing of intron 13 in this transcript is decreased when the neighbouring R/G site, located at position -2 with respect to the 5´ splice site, is edited (Bratt and Ohman, 2003). This effect has been shown to depend solely on the presence of inosine at this position, but not on the steric hindrance for the spliceosome assembly that ADAR binding might cause (Kallman et al., 2003). Editing at the R/G site is also important for the correct alternative splicing of the mutually exclusive downstream exons 14 and 15. Erroneously spliced products where either both alternatively spliced exons are included in or completely omitted from the mature mRNAs have exhibited reduced editing extent at the R/G site, indicating the importance of this editing event for the downstream processing. It is believed that RNA editing at the R/G site only transiently slows down the activation of the adjacent 5´ splice site in order to recruit the branchpoint site either in intron 13 or in intron 14, leading to inclusion of either exon 14 or exon 15 in the mature transcript, respectively (Schoft et al., 2007).

The proximity of the R/G site inosine to the actual 5´ splice site of intron 13 could cause destabilisation in the base-pairing with U1 snRNA and lead to decreased splicing because the wild type adenosine at this position resembles 5´ splice site consensus sequence more closely than inosine introduced upon RNA editing (Mount, 1982). Additionally, an 8 nucleotide stretch surrounding this inosine exhibits high sequence homology to a previously defined splice silencer consensus element (sequence surrounding the R/G site: UAIGgugg, with the intronic sequence in lower case, versus splice silencer consensus: CAGGGUGG) bound by hnRNP A1 or hnRNP H (Fairbrother and Chasin, 2000; Grabowski, 2004).

In this thesis, an effort was made to elucidate which of these two scenarios, if not both, contribute to the decelerated kinetics in the removal of intron 13 of GluR-B pre-mRNA upon RNA editing of the adjacent R/G site.

Part II: SINE RNAs in RNA interference

1.6 SINE RNA interaction with dsRBPs in plants

In the final part of this thesis, the binding affinity of plant SINE-derived RNA intermediates for different double-stranded RNA-binding proteins (dsRBPs) was examined more closely in order to establish the role of these widely distributed RNA retroelements in the regulation of RNA interference.

1.6.1 SINE RNAs

Short interspersed elements (SINEs) are found in genomic sequences of most eukaryotic species. These 80 to 500 basepairs long repetitive regions are evolutionarily related to and descend mostly from tRNAs or, in some cases, from 7SL RNA or 5S RNA. The copy number of SINE elements in eukaryotic genomes ranges from several hundred to several thousand, with the exception of mammals where it can rise up to more than one million copies, as is the case with the human Alu family (Sun et al., 2007). Being non-autonomous RNA polymerase III transcripts, their propagation in the genome, mediated through reverse transcription and the subsequent integration, relies upon their interaction with the enzymatic machinery assembled from the translation products of another widespread group of eukaryotic transposable elements, namely the long interspersed elements (LINEs) (Arnaud et al., 2001; Dewannieux et al., 2003; Ohshima and Okada, 2005).

Several studies have shed light on the role of SINE RNAs in the regulation of gene expression. In mammals, the active transcription of the members of the SINE B2 family is stimulated upon heat shock and results in the repression of general RNA polymerase II-mediated transcription (Allen et al., 2004; Espinoza et al., 2004). Increased transcription of other SINEs, like Alu RNAs, has also been reported upon stress activation. Here, Alu RNAs can indirectly repress translation by controlling the activity of the Protein Kinase R (PKR), a double-stranded RNA-binding protein known for its role in translational repression after stress induction (Schmid, 1998). Alu RNAs, however, can modulate the translational repression by a PKR-independent mechanism, as well (Hasler and Strub, 2006; Rubin et al., 2002). Furthermore, SINE-related BC transcripts in mammalian neurons guide the RNA-binding protein FMRP

to a specific subset of neuronal mRNAs, thereby regulating their translation (Zalfa et al., 2003). Altogether, SINE RNA mechanism of the gene expression control is highly reminiscent of those employed by other PolIII-dependent riboregulators.

The role of SINEs in plants, on the other hand, has remained somewhat obscure. In *Arabidopsis thaliana*, there are six different families of SINEs, located mostly in euchromatic regions and closely associated with genes (Lenoir et al., 2001). Introduction of the transgenic SINE founder locus (SB1) into the genome of *A. thaliana* under the control of its natural promoter lead to a synthesis of the cytoplasmic polyA-tailed SINE transcripts, with a small fraction of transcripts lacking the polyA-tail (Pelissier et al., 2004). The constitutive production of SINE RNAs in these transgenic lines, however, lead to a phenotype exhibiting severe developmental defects, strikingly similar to the phenotype of the plants mutated in the miRNA or trans-acting small interfering (tasi) RNA pathways (Pouch-Pelissier et al., 2008). Moreover, this particular SINE RNA (SB1) belongs to the family of tRNA-derived SINEs. Unlike 7SL-derived SINE RNAs, tRNA-derived ones do not conserve the ancestral RNA folding pattern, but adopt a double-stranded structure consisting of three stem-loops with numerous mismatches and bulges (Sun et al., 2007). This configuration would make them ideal binding partners of dsRBPs, many of which are involved in RNAi-related pathways.

1.6.2 Plant dsRBPs involved in RNA interference

Most of the 17 dsRBPs encoded in the genome of *A. thaliana* are crucial factors in one of the RNAi-related pathways. This family comprises four Dicer-like proteins (DCL1 to DCL4), five double-stranded RNA-binding proteins (HYL1/DRB1 and DRB2 to DRB5) and the HUA Enhancer1 (HEN1) protein (Hiraguri et al., 2005). Such wide spectrum of dsRBD-containing RNAi-related factors in plants is attributed to the specific roles that need to be performed in various RNAi pathways in plants, including miRNA, tasiRNA, viral siRNA, or heterochromatinising siRNA production (Curtin et al., 2008).

HYL1/DRB1 (Hyponastic Leaves 1) is a dsRBD-containing protein in plants shown to be essential for miRNA production, but not for posttranscriptional gene silencing (PTGS) (Vazquez et al., 2004). It is a nuclear protein regulating plant response to many hormones and it is also controlling leaf morphology by cell division,

elongation and polarity (Lu and Fedoroff, 2000). Although HYL1 bears strong homology to *C. elegans* RDE-4 and to *D. melanogaster* R2D2, they are not functionally similar (Vazquez et al., 2004). As previously mentioned, HYL1 interacts with DCL1 in plants in order to produce mature miRNAs, but plays no role in tasiRNA synthesis. This indicates that it preferentially binds dsRNAs containing mismatches and loops, but not perfect double-stranded structures.

DRB4, on the other hand, acts with DCL4 to produce mature tasiRNAs and has no role in miRNA biogenesis. During tasiRNA biogenesis, RNAs transcribed from non–protein-coding *TAS* genes are cleaved by a miRNA–programmed RNA-induced silencing complex. In contrast to classical miRNA targets, RNA-dependent RNA polymerase 6 (RDR6) converts one of the *TAS* RNA cleavage products into double-stranded RNA, which is subsequently processed by DCL4/DRB4 complex to generate a 21-nucleotide tasiRNA population (Nakazawa et al., 2007; Vaucheret, 2005). This suggests that the substrate specificity of DRB4 differs to a great extent from the one of HYL1/DRB1.

2 SPECIFIC AIMS

This thesis consists of three separate topics, each one dealing with different aspects of the function of double-stranded RNA-binding proteins, their association with the respective interaction partners among RNAs and the biological consequences of this macromolecular interplay.

1. Investigating the influence of ADAR-mediated RNA editing near the 5´ splice site on the spliceosome assembly

The R/G site of the GluR-B pre-mRNA found just one nucleotide upstream of intron 13 5´ splice site is a known ADAR target. Upon RNA editing, the splicing efficiency of the neighbouring intron is decreased. R/G site editing destroys the 5´ splice site consensus that base-pairs with U1 snRNA by introducing inosine at the position of adenosine. Furthermore, it creates a sequence similar to splice silencers bound by hnRNP A1 or hnRNP H. Here, we investigated the role of U1 snRNA in the reduced splicing of GluR-B intron 13 after R/G site editing, as well as the association of known splicing repressors with this pre-mRNA.

2. Detecting ADAR1 interaction with microRNA precursors in nucleus and cytoplasm

Both ADAR1 and microRNA precursors are export-competent molecules that interact with Exportin5. Since both nuclear primary microRNAs and cytoplasmic pre-microRNAs were reported as editing targets, we were interested to see if ADAR1, being a shuttling protein, accompanies its microRNA substrates on their Exportin5-mediated nuclear export pathway.

3. Determining SINE RNA binding partners among RNAi-related dsRBPs

Transgenic plants overexpressing SINE RNAs, abundant highly structured retroelements consisting of three stem-loops with mismatches and bulges, have phenotypes similar to RNAi mutants. It is, therefore, plausible that these RNAs, being structurally similar to microRNA intermediates, serve as regulators of RNAi pathways in plants. We investigated *in vitro* interactions of SINE RNAs with different subtypes of dsRBPs involved in one of the RNAi processes to see if the protein affinity for this group of imperfect dsRNAs would match their *in vivo* preferences.

3 MATERIALS AND METHODS

3.1 *In vivo* splicing assay

The system to analyse splicing efficiencies of reporter constructs bearing R/G site in different editing states was used to test the U1 snRNA base-pairing interaction with R/G site upon RNA editing by means of co-transfecting the vectors carrying mutant U1 snRNA genes.

3.1.1 R/G site reporter constructs

Wild type, pre-edited and uneditable R/G site reporter constructs were cloned previously by Vera Schoft in our lab and are described in detail (Schoft et al., 2007), as well as in the Results section of this thesis.

Additionally, the RT-PCR experiment, where the decreased splicing efficiency of the pre-edited R/G site construct when compared to the wild type R/G site constructs was verified, has been performed on total RNA isolated from transfected culture cells. For the first strand cDNA synthesis, oligodeoxyribonucleotide MJ1000a was used (5´ CCTCTACAAATGTGGTATGGCTG 3´). For the amplification, primers MJ1136 (5´ GGTGGAGTTCAAGTCCATCTACATGG 3´; 5´ primer that hybridises to the RFP coding sequence) and MJ1020 (5´ TCGACCAGGATGGGCACCAC 3´; 3´ primer that hybridises to the GFP coding sequence). The corresponding experimental procedures are thoroughly explained below.

3.1.2 U1 snRNA expression vectors

Original vectors for U1 snRNA overexpression were kindly provided by Heiner Schaal from the Institute of Virology, University of Düsseldorf, Germany (Figure 3.1) and were described previously (Murphy et al., 1982). Both wild type and compensatory U1 snRNA mutations were introduced by using mutagenic primers for PCR on the template vector pUCB U1 5A (Figure 3.1 B) and inserting the resulting PCR products into the BglII and XhoI restriction sites of the vector pUCB del U1 (Figure 3.1 A). 5´ mutagenic primers used for this PCR were MJ2392 (5´ CGAAGATCTCATACTTACCTGGCAGGGGAGATACC 3´; BglII restriction site is underlined; mutagenic nucleotide restoring the wild type sequence is highlighted in orange) and MJ2393 (5´ CGAAGATCTCATACTTACCCGGCAGGGGAGATACC 3´; BglII restriction site is underlined; mutagenic nucleotides restoring the wild type

sequence and introducing the compensatory mutation are highlighted in orange and red, respectively) for the wild type and compensatory U1 snRNA mutants, respectively. For both vectors, universal 3´ primer MJ2391 was used in mutagenic PCR reactions (5´ ATCCTCGAGCCTCCACTGTAGGATTAAC 3´; XhoI restriction site is underlined). The resulting PCR products were run on standard 1 % agarose gels, cut out and subjected to double restriction digestion using BglII and XhoI restriction endonucleases (Fermentas). The restricted fragments were ligated into the pre-restricted pUCB del U1. The ligation products were transformed and propagated in *Escherichia coli* strain XL1-Blue. After the isolation and purification of plasmid DNA by the standard alkaline lysis method, the mutated clones were verified by sequencing.

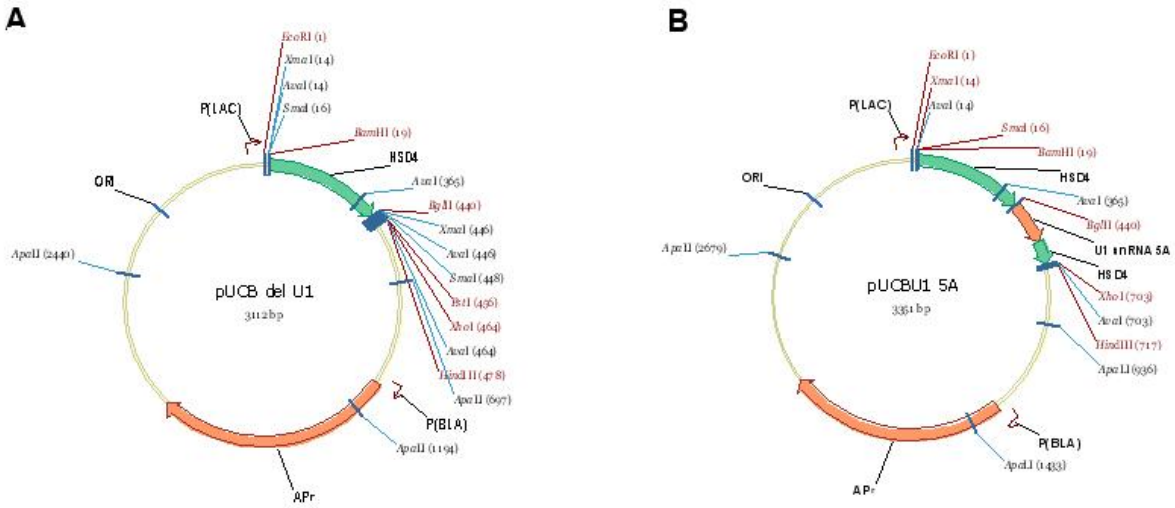

Figure 3.1: Plasmid maps of original U1 snRNA expression vectors

Two depicted plasmids were used to clone wild type and compensatory U1 snRNA expression vectors. In the case of pUCB del U1 (panel A), the U1 snRNA coding sequence together with the 3´ HSD4 sequence was deleted from the HSD4 U1 snRNA expression cassette and this vector was subsequently used to clone the mutagenic PCR products into its BglII/XhoI restriction sites. pUCB U1 5A, on the other hand, was employed as a PCR template to introduce wild type and compensatory mutations into the U1 snRNA coding sequence (panel B).

3.1.3 Primer extension assay

The expression of mutant U1 snRNAs was experimentally validated through primer extension assay. This experimental procedure was applied on total RNAs isolated from HEK293 cells in culture transfected with either wild type or compensatory U1 snRNA expression vectors.

In more detail, a synthetic oligodeoxyribonucleotide primer MJ2413 (5´ CATGGTATCTCCCCTG 3´) complementary to nucleotides 13-28 of U1 snRNA was used to test for the expression of mutant U1 snRNA. This 16-mer was 5´ end-[γ-^{32}P]-labelled with 10 µCi of Redivue [γ-^{32}P] ATP (GE Healthcare) using T4 polynucleotide kinase (New England Biolabs), according to the manufacturer´s instructions. The mixtures of increasing amounts of total RNA and 20 fmol of 5´-radioactively labelled primer were washed, dried and resuspended in the hybridisation buffer (25 mM Tris pH 8.4, 30 mM KCl and 20 mM $MgCl_2$). After brief denaturation (2` on 95ºC), hybridisation was performed for 30 minutes on 42ºC. Extension itself was done in the presence of 0.2 mM dideoxyATP (ddATP), using avian myeloblastosis virus reverse transcriptase (AMV-RT) from Promega, following the company´s protocol. In the presence of ddATP, dCTP, dGTP and dTTP, AMV-RT is capable of adding only 3 nucleotides if the template is wild type U1 snRNA, but adds 7 nucleotides to the compensatory U1 snRNA mutant template. After 40 minutes on 42ºC, the reactions were stopped by ethanol precipitation. The reaction products were analysed on 15 % 8M urea-denaturing 19:1 polyacrylamide gels. After the run, the gels were fixed in a fixing solution (10% methanol, 10% acidic acid), dried for 3 hours and exposed to Kodak PhosphorImager screens. After varying exposure times, the signals were read out, analysed and quantified with the QuantityOne software (Biorad).

3.1.4 Cell culture and transfection

Human embryonic kidney cells (HEK293) used for co-transfection experiments were cultured at 37ºC in Dulbecco´s modified Eagle medium (DMEM) supplemented with 10 % FCS, 0.2 mM L-glutamine and antibiotics. They were incubated in a humid atmosphere with 5 % CO_2.

Cells were transiently transfected using the Nanofectin transfection reagent (PAA) according to the manufacturer´s instructions and FACS-analysed 24 hours after the final transfection round.

3.1.5 FACS analysis

Free access to FACScalibur system (Becton & Dickinson) was kindly provided by Gabriele Stengl from the Biooptics Department on the Institute of Molecular Pathology, Vienna, Austria. The CellQuest 3.3 software was used to collect the flow cytometry data, which was then further analysed with the FlowJo 8.3.1 software. The subsequent statistical analyses were performed in Microsoft Office Excel 2007.

3.2 Co-immunoprecipitation and related procedures

The co-immunoprecipitation approach in this study was exerted to identify RNA-protein interactions between mRNAs derived from R/G site reporter constructs and various hnRNPs on one hand, and microRNA precursors and ADAR1 on the other hand. In the former case, HEK293 cell line was used for transfection and subsequent co-immunoprecipitation, while in the latter case the cell line of choice was HeLa.

3.2.1 UV crosslinking

In order to stabilise the interaction of mRNAs derived from R/G site reporter constructs with hnRNPs A1 and H, UV crosslinking was performed using UV Stratalinker 2400 (Stratagene) as previously described (Dreyfuss et al., 1984).

3.2.2 Subcellular fractionation

In the second part of this study, we wanted to determine whether different ADAR1 isoforms interact with microRNA precursors in both nucleus and the cytoplasm. This is why these two subcellular compartments had to be separated. Immediately after gentle harvesting and washing, HeLa cell pellet was resuspended in 5 volumes of precooled Earle´s balanced salt solution (0.15 M NaCl, 16 mM K_2HPO_4, 2.5 mM KH_2PO_4 at pH 7.4), washed, pelleted and again resuspended, this time in 10 volumes of hypotonic RSB solution (10 mM NaCl, 1.5 mM $MgCl_2$ and 10 mM Tris pH 7.4) and left on ice. Then, the swollen cells were homogenised by 10-12 tight-fitting pestle strokes in Dounce glass homogeniser. Once the homogenised cell suspension had a free nuclei:intact cell ratio of 9:1, it was centrifuged to get the cytoplasmic fraction in the supernatant and the nuclear fraction in the pellet.

3.2.3 Co-immunoprecipitation

Whole cell, cytoplasmic or nuclear fractions were sonicated and precleared in NET-2 buffer (150 mM NaCl, 50 mM Tris pH 7.4 and 0.05 % NP-40) to obtain lysates for immunoprecipitations. The corresponding primary antibodies (mouse 10D1/α-hnRNP A1 for hnRNP A1 at 1:200; rabbit α-hnRNP H from Abcam for hnRNP H at 1:200; rabbit Sat309 for both ADAR1 isoforms at 1:50; rabbit Sat1146 for ADAR1p150 at 1:50) were coupled to Protein A Sepharose beads (Amersham Biosciences) by overnight incubation on 4ºC. Then, the precleared lysates were added to the coupled antibodies in overall volume of 500 µl and immunoprecipitation was allowed to take place for 1 h on 4ºC. After this, the supernatant was removed and the beads were washed 5 times with NET-2 buffer. Then, 1/10 of beads was immediately resuspended in equal volume of 2xSDS sample buffer, boiled for 2 minutes and used for the protein detection via Western Blot, while the rest was subjected to the co-immunoprecipitated RNA isolation, followed by RT-PCR.

3.2.4 SDS polyacrylamide gel electrophoresis

Protein separation on the molecular weight basis is achieved in SDS polyacrylamide gels. Such typical gel consists of resolving part (? % 30:1 acrylamide:bis-acrylamide, 375 mM Tris pH 8.8 and 0.1 % SDS) and a stacking part (used to concentrate negatively charged proteins to a thin boundary; 3.6 % 30:1 acrylamide:bis-acrylamide, 110 mM Tris pH 6.8, 0.1 % SDS and 13 % glycerol). The proteins boiled in equal volume of 2xSDS sample buffer (220 mM Tris pH 6.8, 2 % SDS, 20 % glycerol, 0.03 % bromophenol-blue and 2 % β-mercaptoethanol) were loaded onto such gel. Electrophoresis was performed in 1xSDS running buffer (3 g/l Tris-HCl, 14.2 g/l glycine and 1 g/l SDS) at 20 mA constant electric current.

3.2.5 Western Blotting

The electrophoresed proteins were transferred from the gel to the nitrocellulose membrane by means of tank blotting. For this purpose, the equilibrated membrane was placed on a gel between Whatman papers and sponges. The resulting blotting sandwich was put in an icepack-containing tank filled with Transfer

buffer (20 mM Tris-HCl and 150 mM glycine) and subjected to transfer for 1 hour at 360 mA.

To visualise protein marker, after blotting the membrane was incubated in Ponceau S solution for 10 minutes and washed with water until protein bands became visible. The membrane was then blocked in 5 % dry milk in 1xTBS-T (Tris-buffered saline with 0.05 % Tween-20) for 10´ at room temperature. After that, the primary antibody solutions in 5 % dry milk in 1xTBS-T (rabbit α-hnRNP H at 1:10000; rabbit α-CBP80 at 1:1000; mouse 10D1/α-hnRNP A1 at 1:1000; rat Sat12/α-ADAR1 at 1:1000) were applied to the membrane and shaken for 1 hour. Later on, after 3 washing steps of 5 minutes in 0.5 % dry milk in 1xTBS-T, the membrane was incubated in one of the secondary antibody solutions in 5 % dry milk in 1xTBS-T (goat alkaline phosphatase-conjugated α-rabbit antibody at 1:1000; goat alkaline phosphatase-conjugated α-mouse antibody at 1:1000; goat horseraddish peroxidase-conjugated α-rat antibody at 1:20000). This was also followed by three washing steps, each 5 minutes long.

Blotted proteins were then detected either by chemiluminescence or by NBT/BCIP chromogen precipitation. If the secondary antibody was coupled to horseraddish peroxidase, the chemiluminescence was induced and detected by using SuperSignal Solution (Pierce) according to the manufacturer´s instructions. On the other hand, the secondary antibodies coupled to alkaline phosphatase require a chromogenic substrate NBT/BCIP (nitro-blue tetrazolium/5-bromo-4-chloro-3-indolylphosphate-p-toluidine), which in the presence of this enzyme forms a dark precipitate on the membrane.

3.2.6 RNA isolation and RT-PCR

RNA isolation was performed by either directling applying the RNA isolation reagent Trizol (Invitrogen) to cultured cells in dishes or applying it to cell lysates or Protein A Sepharose beads with co-immunoprecipitated RNAs, all using the conditions recommended by the manufacturer.

Two consecutive DNaseI (Roche) digestion steps were necessary to get rid of any traces of residual DNA which might lead to false positive signals in PCR and were made according to the manufacturer´s protocol.

Purified RNAs were then reversely transcribed with RevertAid M-MuLV Reverse Transcriptase (Fermentas) following the company´s instructions, using either MJ1000a for R/G site construct mRNAs or random hexanucleotide mixture (Roche) for priming miRNAs and γ-tubulin B-subunit mRNA.

1/10 of heat-inactivated RT reactions served as a template for PCR reaction. Respective plasmids (in the case of transfections) and cell lysates (in all cases) were employed as control templates.

Following primer combinations were used in PCRs:

MJ1136 (5´ GGTGGAGTTCAAGTCCATCTACATGG 3´) and MJ1020 (5´ TCGACCAGGATGGGCACCAC 3´) for all the amplifications of R/G site constructs-derived cDNAs;

MJ1515 (5´ GGCAGAGGGCAACAGTTCTTC 3´) and MJ1516 (5´ GGCTGAGCCGCAGTAGTTCTTC 3´) for the amplification of cDNA derived from the miR-22 precursor;

MJ1800 (5´ TGTCAGACAGCCCATCGACTG 3´) and MJ1801 (5´ TGTCGGGTAGCTTATCAGACTG 3´) for the amplification of cDNA derived from the miR-21 precursor;

MJ1025 (5´ GACAACACAGCCCTGAACCG 3´) and MJ1026 (5´ GACTGGTCTGTAGTGAGCGG 3´) for the amplification of cDNA derived from the γ-tubulin B-subunit mRNA.

The PCR products were then analysed on 1-2 % agarose gels in either 1xTAE or 1xTBE running buffers.

3.3 Nuclease protection assay

To map regions of SINE RNAs that interact with different dsRBD-containing proteins, RNase V1 protection assay was employed.

First, RNA structures were predicted using RNA/DNA folding and hybridisation software Mfold, version 2.3 (Zuker, 2003).

Then, RNA was *in vitro* transcribed and radioactively trace-labelled (for quantification purposes) from linearised T7 promoter-containing vector using recombinant T7 RNA polymerase and [α-^{32}P] ATP (GE Healthcare). Transcripts were gel-purified and ethanol-precipitated.

2 pmol of RNA was then dephosphorylated using calf intestinal phosphatase (New England Biolabs) according to manufacturer's protocol. Dephosphorylated RNAs were 5′-end labelled with T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P] ATP (GE Healthcare).

Gel-purified 5′-labeled RNAs were subsequently used for nuclease protection assays (Hallegger et al., 2006). RNase V1 recognizes any 4-6-nt segment of polynucleotide backbone with an approximately helical conformation and cleaves leaving 5′-phosphate (Lowman and Draper, 1986). For the partial digest with RNase V1, 20 fmol (corresponding to 50,000 cpm) of RNA were centrifuged, washed, dried and resuspended in structure buffer (Ambion: 100 mM Tris at pH 7, 1 M KCl, 100 mM $MgCl_2$). After annealing, 1 µL of tRNA (1 µg/µL, Ambion) was added, followed by the addition of increasing protein concentrations (50 nM, 150 nM and 500 nM). To ensure protein binding to RNA, samples were incubated for 15 min at room temperature. Then, 0.005 units of RNase V1 (Ambion) were added and the reactions were incubated for additional 10 min at room temperature. Reactions were stopped by ethanol/salt precipitation. Samples were loaded together with alkaline hydrolysis ladder and denaturing RNase A (Ambion) and RNase T1 (Boehringer Mannheim) digests of RNAs on thin denaturing RNA gels.

After the run, the gels were dried for 3-4 hours and exposed to Kodak PhosphorImager screens. After varying exposure times, the signals were read out, analysed and quantified with the QuantityOne software (Biorad).

4 RESULTS

4.1 Molecular mechanism of altered pre-mRNA splicing upon RNA editing

A to I RNA editing is a posttranscriptional modification that contributes significantly to transcriptome diversification. Occurrence of editing sites in exons of mRNAs, however, has the potential of enriching the proteome repertoire, as well. In fact, the most prominent coding targets of ADARs identified thus far are messages coding for various receptors in the mammalian central nervous system: subunits of glutamate receptor superfamily, serotonin 5-HT_{2C}-receptor and potassium channel KCNA1. In all of them, amino acid substitutions due to editing have major impact on properties of given proteins. Recently, a bioinformatics screen was used to identify novel editing targets and the substrates were verified experimentally (filamin A, CYFIP2, BLCAP and IGFBP7 mRNAs) (Levanon et al., 2005). Additionally, a genome-wide identification by parallel DNA capturing and sequencing greatly expanded the number of verified targets (Li et al., 2009). Interestingly, in some targets the editing sites are found in close proximity to 5´ splice sites where exon-intron base-pairing interactions are formed to provide a suitable binding platform for ADARs. Editing in these positions can have a profound effect on initial steps of the splicing machinery recruitment, e.g. U1 snRNA base-pairing to exon-intron junction.

Previously, a cell-based assay to monitor splicing efficiency established in our lab was employed to show effects of editing of R/G and Q/R sites on splicing in GluR-B mRNA. Results from this screen prompted us to apply this method to take a closer look on the efficiency of U1 snRNP recruitment to edited 5´ splice sites in mRNAs, emphasising the importance of altered U1 snRNA::mRNA interaction potential upon RNA editing.

Furthermore, high resemblance of the edited exon-intron junction regions to splice silencer consensus sequence suggests that the changes in splicing efficiency could be attributed to binding of splicing repressors, like hnRNP A1 or hnRNP H proteins which show high affinity for G-rich sequences.

In this study, both of these possibilities were addressed in detail in order to elucidate the exact step in the pre-mRNA processing pathway that gets affected by adenosine deamination, leading to alterations in the splicing of editing substrates and, thereby, enriching the proteome repertoire in cells and tissues where ADAR family of enzymes exhibits catalytic activity.

4.1.1 Testing the imperfect 5´splice site hypothesis

4.1.1.1 *In vivo* splicing assay

As seen in Figure 4.1, an *in vivo* splicing assay already established in our lab by Vera Schoft was implemented to investigate the influence of the R/G site editing on the U1 snRNA base-pairing affinity for the intron 13 5´ splice site and its subsequent removal. This system was originally developed to monitor the splicing efficiencies of GluR-B pre-mRNA reporter constructs in different editing states. Coding sequences for red (RFP) and green (GFP) fluorescent proteins introduced into a mammalian expression vector were separated by a flexible region of the *E. coli* lacZ gene and a multiple cloning site (MCS) to maintain the open reading frame (ORF). The MCS was then used to insert the genomic sequences containing the exons and intron with the editing site of interest. In all these constructs, RFP is expressed constitutively, while the expression of GFP as a downstream reporter relies on the elimination of the stop codon-containing intron, splicing of which is severely influenced by RNA editing. Another special feature of the aforementioned constructs was the nuclear localisation signal (NLS) to allow for an unequivocal quantification of fluorescence during microscopic and FACS analysis.

For this method, R/G site constructs in three different “editing states” were used (Figure 4.1 B): the wild-type construct (WT), the construct with the R/G site adenosine mutated to guanosine to mimic inosine (Pre-edited) and, finally, the construct with mutated editing complementary sequence (ECS), preventing the formation of the double-stranded structure and consequently ADAR binding (Uneditable). It is important to have the “uneditable” construct as an internal control because of the background level editing of the “wild-type” construct in the biological system used for the experiment, i.e. HEK293 cells.

In this part of the study, the U1 snRNA base-pairing potential upon RNA editing was investigated by co-expressing U1 snRNA mutants with reporter constructs in human cells in culture (HEK293).

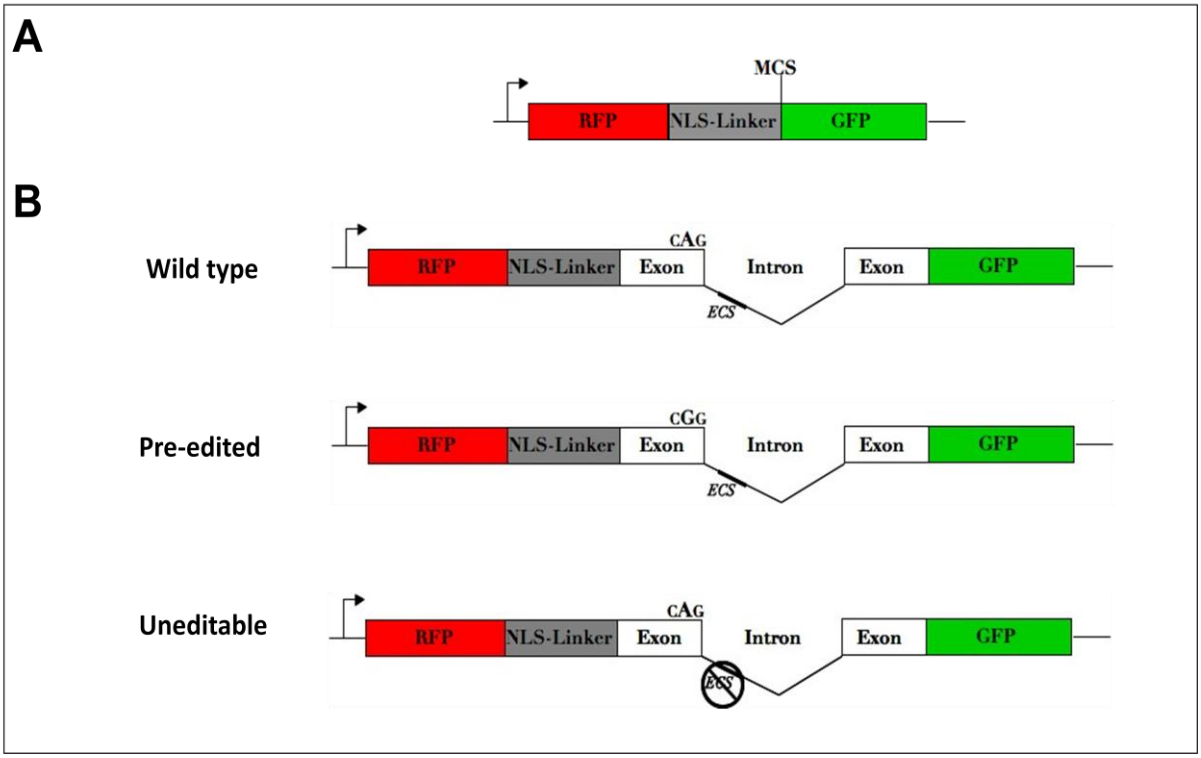

Figure 4.1: *In vivo* splicing assay

A: mammalian expression vector with MCS between RFP- and GFP-coding sequences enables insertion of ADAR substrates and monitoring of splicing efficiencies in different RNA editing backgrounds by means of measuring ratios of green versus red fluorescence

B: Schematic representation of different types of editing constructs used in the study: in pre-edited constructs adenosine (A) was mutated to guanosine (G) to mimic inosine (I), a product of hydrolytic deamination upon RNA editing; uneditable constructs were made by mutating editing complementary site (ECS), thus preventing ADAR binding.

4.1.1.2 Reduction of splicing in pre-edited R/G site construct

In contrast to the fully spliced wild type R/G site construct, there is still a substantial fraction of unspliced pre-edited pre-mRNA present in the transfected cells, as judged by the RT-PCR experiments (Figure 4.2).

The aim of the overexpression of compensatory U1 snRNA mutant together with the pre-edited construct was to facilitate splicing for this splice-deficient substrate and restore the wild type splicing levels.

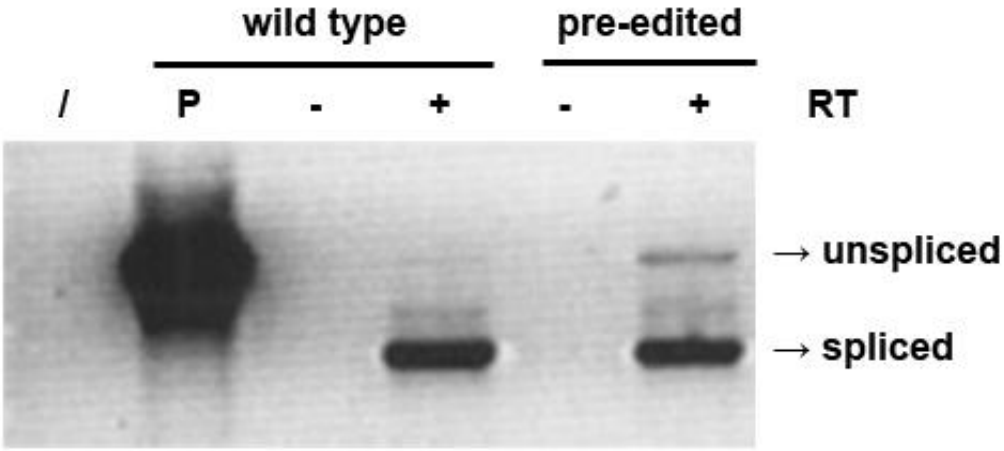

Figure 4.2: Splicing in pre-edited R/G construct is decreased

RT-PCR analysis reveals significant amounts of unspliced pre-edited message (“pre-edited +” lane) when compared to the wild-type mRNA (“wild type +” lane). Wild type plasmid was used as a positive control for unspliced messages (“P” lane).

4.1.1.3 Overexpression of U1 snRNA mutants

The remarkable complementarity between the 5´ splice site consensus sequence and the sequence of U1 snRNA led to the proposal that the nucleotides 2-10 at the 5´ end of U1 snRNA recognise the positions -3 to +6 in the mRNA precursors, with respect to the 5´ splice site. This was further corroborated by the variety of biochemical and evolutionary evidence, such as the accessibility of the 5´ end of U1 snRNA in the intact U1 snRNP, the abolishment of the *in vitro* splicing activity upon RNAse H treatment of U1 snRNP in the presence of the oligonucleotide complementary to the 5´ end of U1 snRNA, or the sequence conservation of the 5´ end of U1 snRNA in species as diverse as human, rat, mouse, chicken, frog, sea urchin or fruit fly (Zhuang and Weiner, 1986). Most importantly, the nucleotide changes in the 5´ splice sites of pre-mRNAs, reducing the match to the U1 snRNA consensus sequence, were usually found to weaken the splicing effectiveness. This scenario, of course, is the one easy to imagine taking place in the case of R/G site adenosine deamination at position -2 with respect to the 5´ splice site (Mount, 1982).

Therefore, the mutant U1 snRNA construct was designed in this study to compensate for the splicing reduction in the editing substrates by means of the site-directed mutagenesis at the nucleotide position base pairing with the R/G site of the GluR-B pre-mRNA (the original pUC-backbone U1 snRNA vectors were kind gift from Heiner Schaal, Institute of Virology, University of Düsseldorf, Germany and were previously described (Murphy et al., 1982).

To verify expression of mutant U1 snRNA, primer extension assay on total RNAs isolated from transfected HEK293 cells was performed. In Figure 4.3 B, a cloverleaf U1 snRNA structure was used to depict 5´ end-labelled primer and its complementary sequence, as well as the uridines marking the reverse transcription stop when using dideoxyATP (ddATP) for wild type and compensatory U1 snRNA transcripts, respectively. Determined from band intensities in Figure 4.3 B, the expression levels of compensatory U1 snRNA mutant are boosted up to 15% of the expression levels of wild type U1 snRNA (Figure 4.3 C).

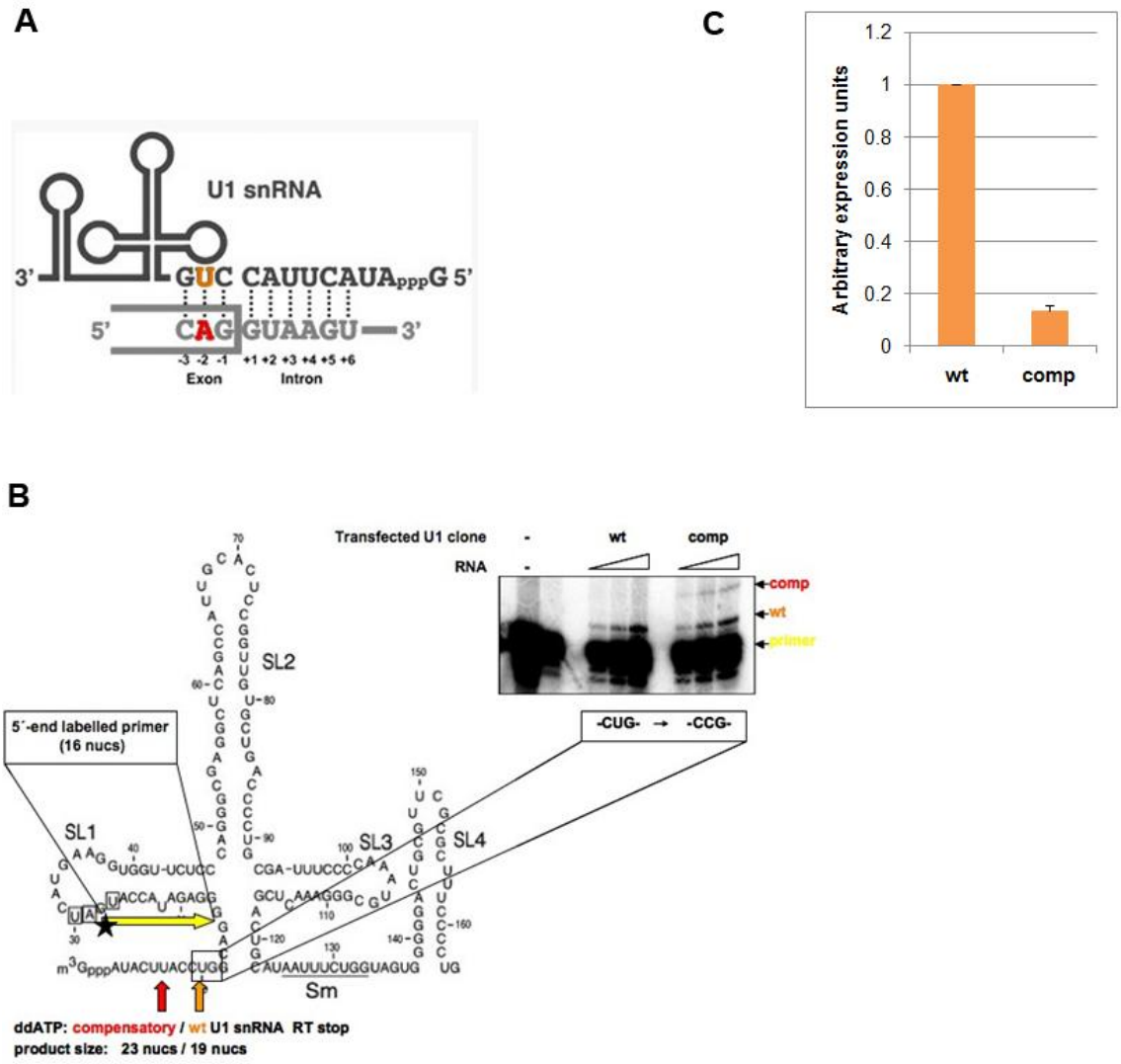

Figure 4.3: U1 snRNA overexpression

Editing in GluR-B pre-mRNA affects the -2 position of the 5´splice site that base-pairs with 5´ end of the U1 snRNA. Therefore, compensatory point mutation was introduced into expression vector bearing U1 snRNA gene at a position base-pairing with inosine/guanosine in edited/pre-edited substrates in order to restore its base-pairing potential.

A: Base-pairing interaction between U1 snRNA and 5´ splice site. The base pair affected by RNA editing is highlighted in red and orange in pre-mRNA and U1 snRNA, respectively.

B: The cloverleaf secondary structure of U1 snRNA with marked complementary sequence for 5´ end-labelled primer (yellow) and ddATP-RT (dideoxyATP) reverse transcription stops for wild type (orange) and compensatory (red) U1 snRNAs. Increasing amounts of total RNA extracts from either wild-type or compensatory U1 snRNA HEK293 transfectants were added to the primer extension reactions. Samples were analysed on gel (upper right panel).

C: U1 snRNA compensatory mutant expression was determined by quantification of band intensities from primer extension assay and it reaches up to 15% of endogenous wild type U1 snRNA expression

Since the number of transfected cells under our transfecting conditions ranges from 20-40% and the level of compensatory U1 snRNA mutant expression reaches 15% of the wild type U1 snRNA levels (Figure 4.3 C), it can easily be estimated that the fraction of compensatory U1 snRNA mutant in cells that have actually been transfected could be as high as 50%.

The confirmation of such a substantial level of mutant expression was indispensable to proceed with the experiments where mRNAs with editing site near the 5´ splice site and compensatory U1 snRNA mutants are brought together into the same biological system. This would enable them to interact intimately, because of their high level of complementarity, especially in comparison with the affinities they both exhibit for their respective “wild type” base-pairing interaction partners.

4.1.1.4 Co-transfection experiments

As already mentioned, the method of choice to investigate the base-pairing potential between 5´splice site of edited mRNAs and U1 snRNA was co-transfection.

The cells were seeded out 12-24 hours at low density before the first round of transfection. When cells have reached 20-30% of confluency, they were first transfected with the increasing amounts of U1 snRNA-bearing constructs to allow for the assembly of U1 snRNPs containing mutant U1 snRNAs. 24 hours after this, the second round of transfection, the one with the R/G site reporter constructs carrying genes coding for RFP and GFP, was performed. On the following day, 24 hours after the second and final transfection round, the cells were harvested and processed for the FACS analysis.

Cell line of choice for co-transfection was HEK293.

4.1.1.5 FACS analysis

Offering the possibility of analysing large cell populations in relatively short amount of time, fluorescence-activated cell sorting (FACS) was employed to collect data for the statistically significant number of cells exhibiting red and green fluorescence. Intensities of red and green fluorescence of around 30 000 cells transfected first with U1 snRNA constructs and then with R/G site reporter constructs were analysed in this manner 24 hours after the second transfection round.

in Figure 4.4, the representative 2D dot plots of different reporter constructs are depicted. The gate chosen for further analysis encompasses cells which exhibit medium red fluorescence (y-axis: FL2-H). The green fluorescence (x-axis: FL1-H) can only be detected after productive splicing occured and its threshold was set to the arbitrary value of 10 for this particular gate. Empty reporter vector (RNLG) was used as a control for constitutive red and green fluorescence (uppermost panel). The wild type R/G constructs (second panel from top) are showing high green fluorescence intensities for the chosen cell population. In the cells transfected with the pre-edited R/G constructs (second panel from bottom), due to impaired splicing, a decrease in green fluorescence is observed (a shift of the cell population to the left along the x-axis): in order to “rescue” splicing of this particular construct (restore the original green fluorescence intensities seen in the wild type construct), a compensatory U1 snRNA mutant was being co-transfected. Cells with the uneditable construct (lowest panel) evade RNA editing and, thus, show higher green fluorescence signals (a shift back to the right along the x-axis).

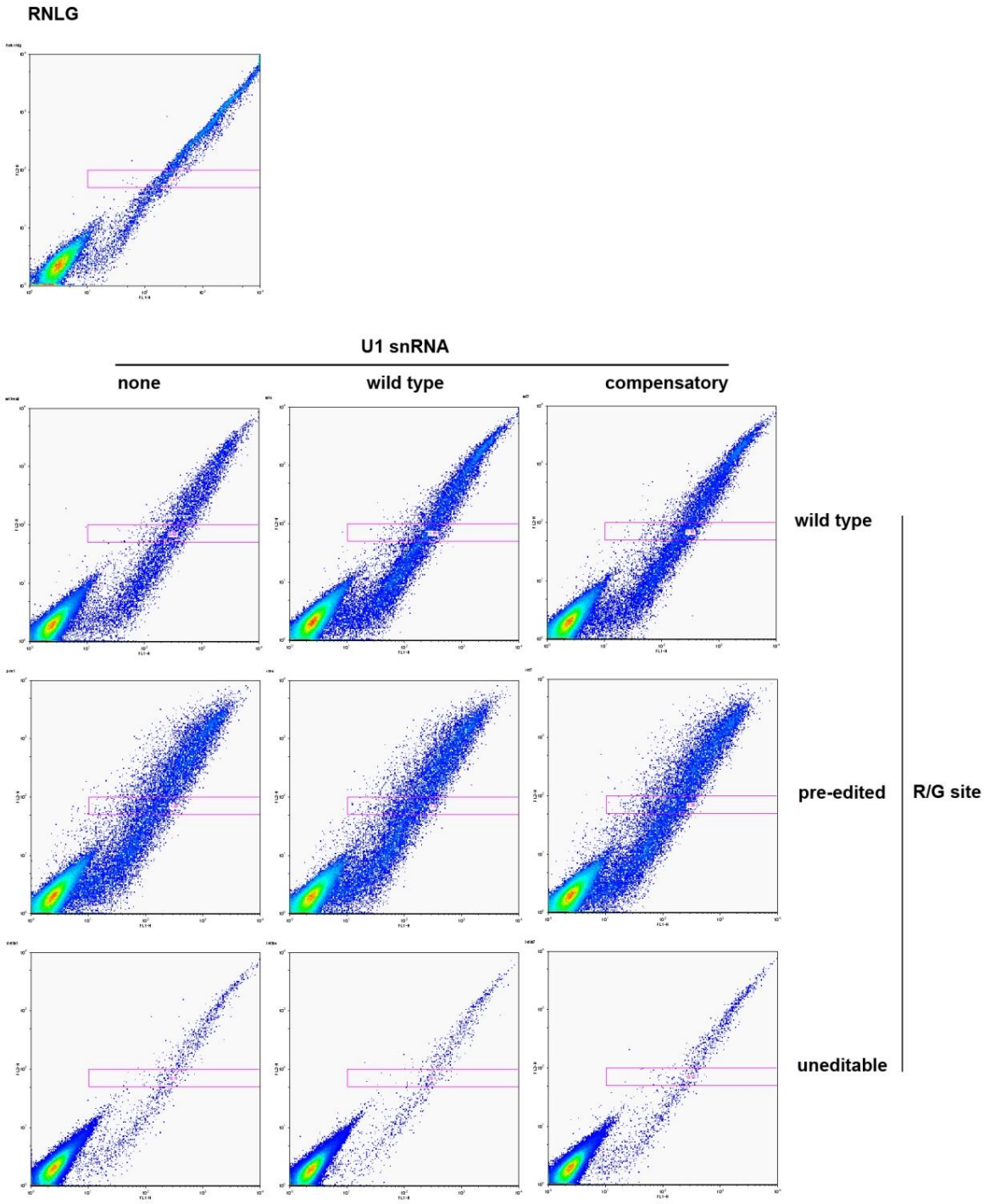

Figure 4.4: FACS analysis of R/G site constructs

2D plots of red (FL2-H; y-axis) versus green fluorescence (FL1-H; x-axis) are depicted for representative HEK293 transfections of control RNLG vector (containing RFP-GFP fusion protein) and wild type, pre-edited and uneditable R/G site constructs, with or without wild type or compensatory U1 snRNA constructs. Gated events chosen for further statistical analysis are shown in purple, the number indicating the percentage of total events falling into this subset.

The statistical analysis of the gated events was performed by first calculating ratios of green versus red fluorescence intensities for each individual cell. Afterwards, the mean value for every sample, i.e. every construct combination could be obtained, together with the values for the standard deviation. The statistical significance of the observed differences was determined through application of paired Student´s T-test with two-tailed equal distribution.

In Figure 4.5, the behaviour of each R/G site construct in combination with either wild type or compensatory U1 snRNA is presented as a bar diagram of mean ratios of green versus red fluorescence signals as the indicator of splicing efficiency, with standard deviation values depicted as error bars in the gated window.

As seen in panel A of Figure 4.5, a slight increase in splicing can be observed for the wild type R/G site construct when the compensatory U1 snRNA construct is pre-transfected (lane 7), compared to the wild type U1 snRNA overexpression (lane 4). This effect can be attributed to the residual ADAR activity in the HEK293 cell line.

In the case of the pre-edited R/G site construct (Figure 4.5, panel B), overexpression of compensatory U1 snRNA mutant leads to a 10% increase in green fluorescence (lane 7), indicating that the partial restoration of the wild type splicing efficiency occurs once the modified -2 position of the 5´splice site is counteracted by the mutation of its U1 snRNA base-pairing partner.

The slight increase (less than 3%) in the splicing efficiency of the wild type R/G site construct upon compensatory U1 snRNA overexpression, when compared to the wild type U1 snRNA overexpression, could be attributed either to the effect that the U1 snRNA vector has on the system or to the residual ADAR activity in HEK293 cells. In the latter case, ADAR would convert some adenosines at the -2 position (with respect to the 5´ splice site) of this construct into inosines, resulting in the decrease of the splicing efficiency. Upon compensatory U1 snRNA overexpression, however, a slight increase in splicing would be observed because the inosine at -2 position is able to base-pair with the compensatory cytidine in U1 snRNA, leading back to more productive splicing. To distinguish between these two scenarios, the uneditable R/G site construct that can not be bound and modified by ADAR was used (Figure 4.5, panel C). As expected, this construct with high splicing efficiency shows a decrease in green fluorescence intensity in combination with the compensatory U1 snRNA (lane 7), when compared to the wild type U1 snRNA co-expression (lane 4) because of its editing-resistant “wild type” 5´ splice site sequence.

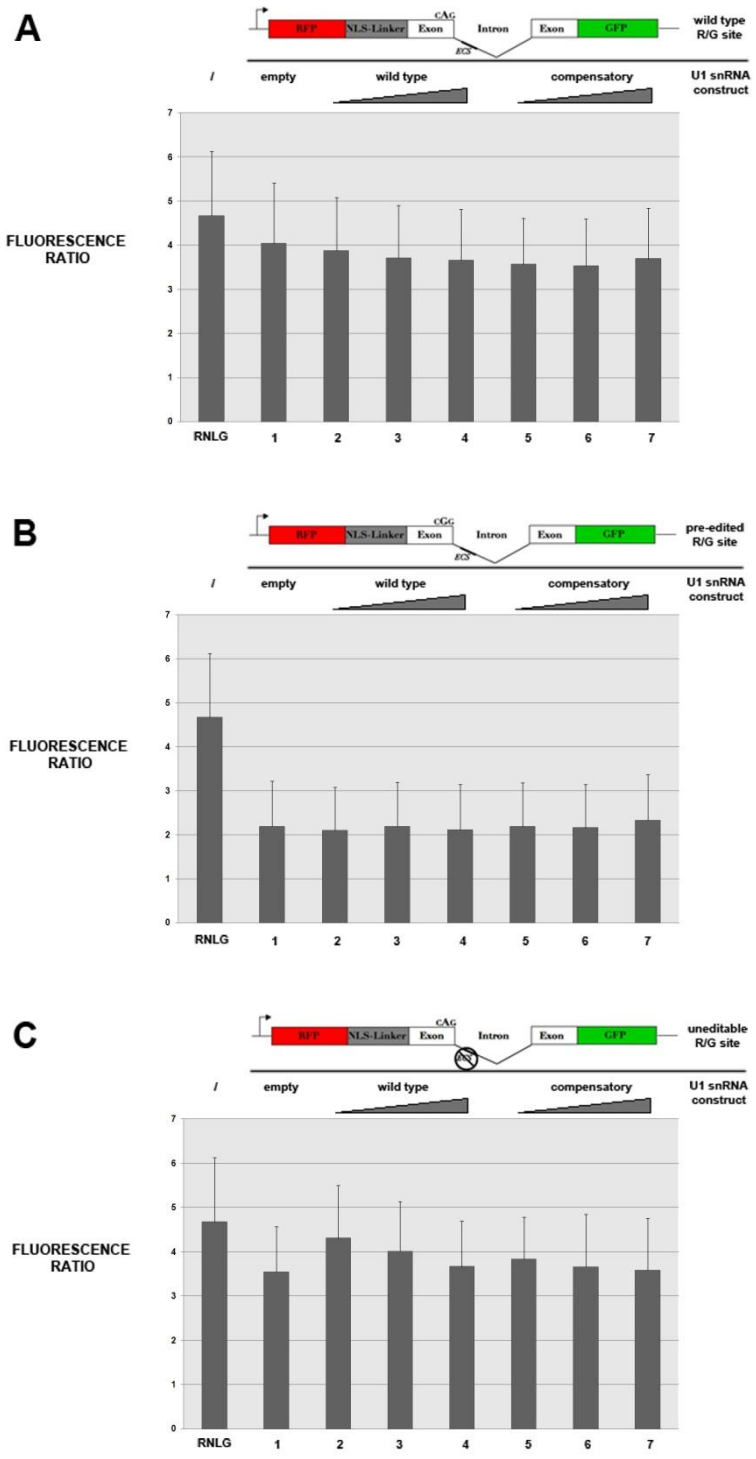

Figure 4.5: Influence of U1 snRNA overexpression on splicing of R/G site constructs

Wild type (panel A), pre-edited (panel B) and uneditable (panel C) R/G site constructs were cotransfected with either empty vector (lanes 1) or increasing amounts of wild type (lanes 2-4) and compensatory U1 snRNA constructs (lanes 5-7). Mean values of green versus red fluorescence intensities for each sample were plotted as bar diagrams. RNLG vector was used as control (lanes RNLG).

In Figure 4.6, the data from the previous figure are summarised by including only the samples where the highest amounts of U1 snRNA constructs were cotransfected: this data set is showing the strongest effect of U1 snRNA overexpression, especially in restoring the wild type splicing “phenotype” for the pre-edited R/G site construct when compensatory U1 snRNA is introduced into the system (Figure 4.6, lane 6). Furthermore, the observed splicing restoration for the aforementioned construct pair (around 10 %) is in line with the values expected for such transfection rates (approximately 15 %). In any case, only a moderate effect on splicing restoration could be seen in wild type R/G construct (Figure 4.6, lane 3), while in the case of uneditable R/G construct no significant change was observed (Figure 4.6, lane 9).

.

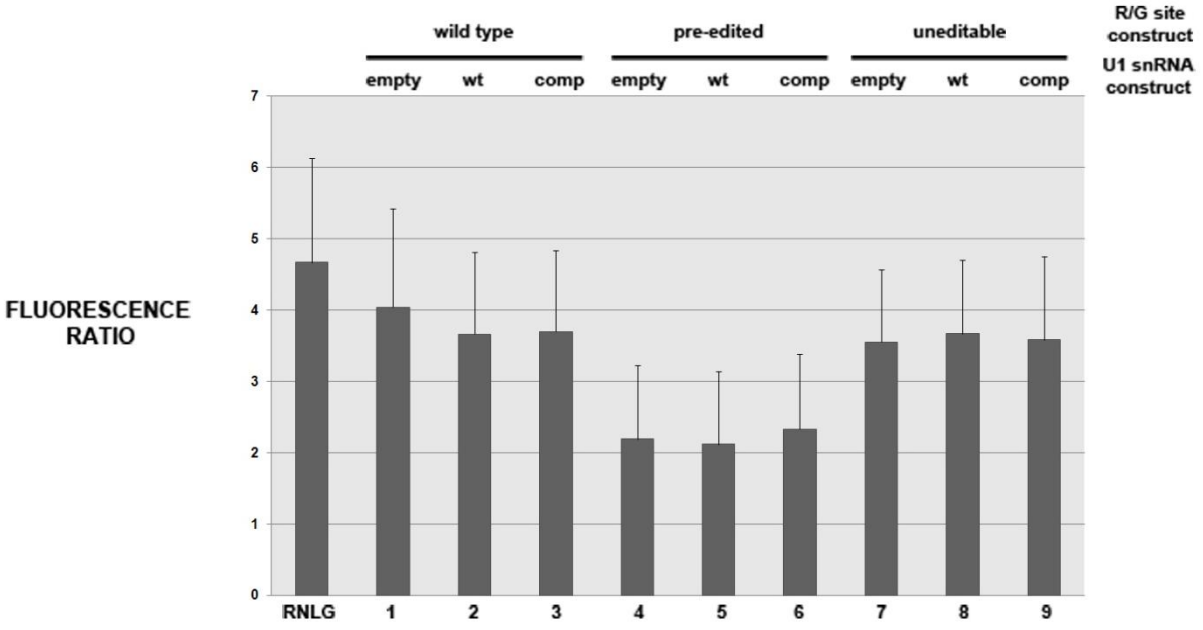

Figure 4.6: Summary of FACS data for the cotransfections of U1 snRNA and R/G site constructs

Bar diagram shows mean values of green versus red fluorescence intensities for R/G site constructs cotransfected with the highest amounts (2 µg) of U1 snRNA constructs. Values for standard deviation are shown as error bars.

The statistical significance of the observed differences in the analysed FACS data was determined by using paired Student´s T-test with two-tailed equal distribution. This statistical application detects the differences for the given parameters in two samples, expresses them as p-values and indicates if they are

high enough to consider the samples dissimilar. In Table 4.1, for each type of R/G site construct the p-values were calculated comparing the values of green versus red fluorescence ratios for every single event (cell) in the two samples being tested.

As expected, the lowest p-values, i.e. the highest difference is derived for the set of pre-edited R/G site constructs, with compensatory U1 snRNA cotransfectant showing most notable discrepancy compared to both empty and wild type U1 snRNA ones. The differences are lowest for the set of wild type R/G site cotransfectants, showing that the U1 snRNA mutants influence the splicing of this pre-mRNA the least.

These data go in line with the previous results and serve as an additional confirmation for the role of compensatory U1 snRNA mutant in reconstituting to some extent the splicing efficiency of the pre-edited R/G site construct.

Table 4.1: Student´s T-test analysis of observed differences in FACS data

Three main rows labelled on the left side (wild type, pre-edited and uneditable) indicate the transfected R/G site construct. The co-transfected U1 snRNA constructs (empty, wild type and compensatory) are indicated in the first row and in the last column of the table in order to enter p-values calculated by comparing differences between the cell populations carrying the same R/G site construct, but different U1 snRNA vectors. If the p-value obtained through student´s T-test is lower than 0.05, two samples or populations are considered dissimilar.

R/G site	empty	wild type	compensatory	**U1 snRNA**
wild type	/	***/***	***/***	empty
	0.584	***/***	***/***	wild type
	0.394294	0.216722	***/***	compensatory
pre-edited	***/***	***/***	***/***	empty
	0.50352	***/***	***/***	wild type
	0.009289	0.003184	***/***	compensatory
uneditable	***/***	***/***	***/***	empty
	0.019601	***/***	***/***	wild type
	0.687504	0.021714	***/***	compensatory

4.1.2 Interaction of R/G site pre-mRNAs with hnRNPs

Alternative splicing occurs through the usage of different splicing signals, like alternative splice sites or exonic and intronic splicing enhancers and silencers, and contributes massively to the synthesis of various protein isoforms from a unique transcript.

In the case of R/G site in GluR-B pre-mRNA, proper alternative splicing decisions are stimulated indirectly only upon RNA editing, which first leads to the splicing reduction and transient intron 13 retention, followed by correct processing of downstream introns. The resulting region (UUAIGgugg with intron sequence in lower case) exhibits strong sequence homology to an hnRNP A1-binding splice silencer consensus (GGCAGGGUGG). Moreover, the introduction of guanosine-mimicking inosine into this sequence raises the affinity of another splice silencing regulator, G-rich binding hnRNP H, for this region. Thus, we decided to analyse the differences in association of wild type and pre-edited R/G site pre-mRNAs with candidate hnRNPs.

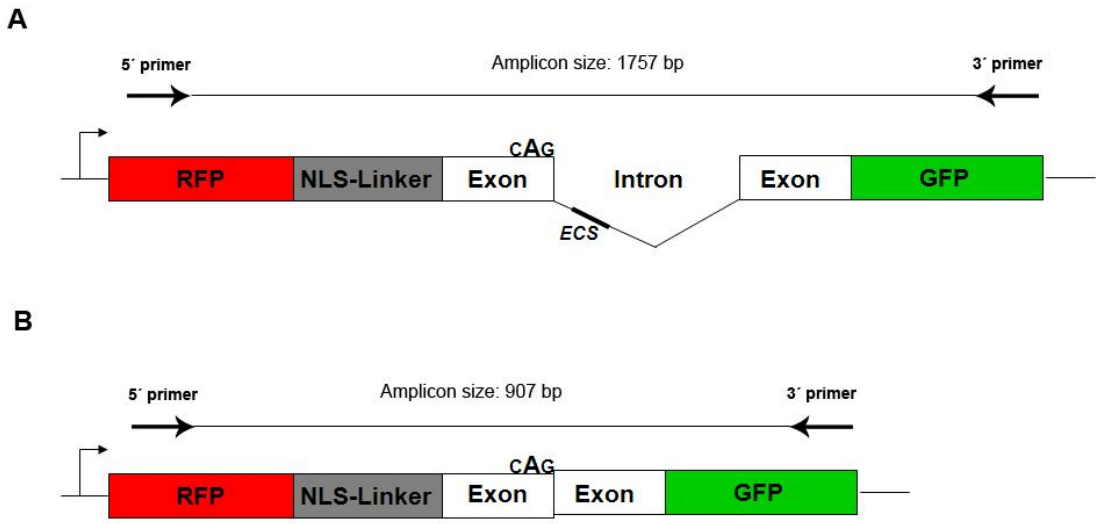

Figure 4.7: Schematic drawing of RT-PCR for R/G site construct pre-mRNAs

RT-PCR products with expected amplicon sizes are depicted for unspliced and spliced versions of R/G site construct messages in panels A and B, respectively.

HEK293 cells were transfected with the reporter constructs and immediately before harvesting UV-crosslinking was performed to stabilise the interaction between mRNAs and hnRNPs (Dreyfuss et al., 1984). Cell lysates were then applied to Protein A Sepharose beads with covalently coupled anti-hnRNP A1 or anti-hnRNP H

antibodies for co-immunoprecipitation. The co-immunoprecipitated RNAs were further investigated by RT-PCR, using exonic primers that allow for unambiguous distinction between spliced and unspliced products (Figure 4.7).

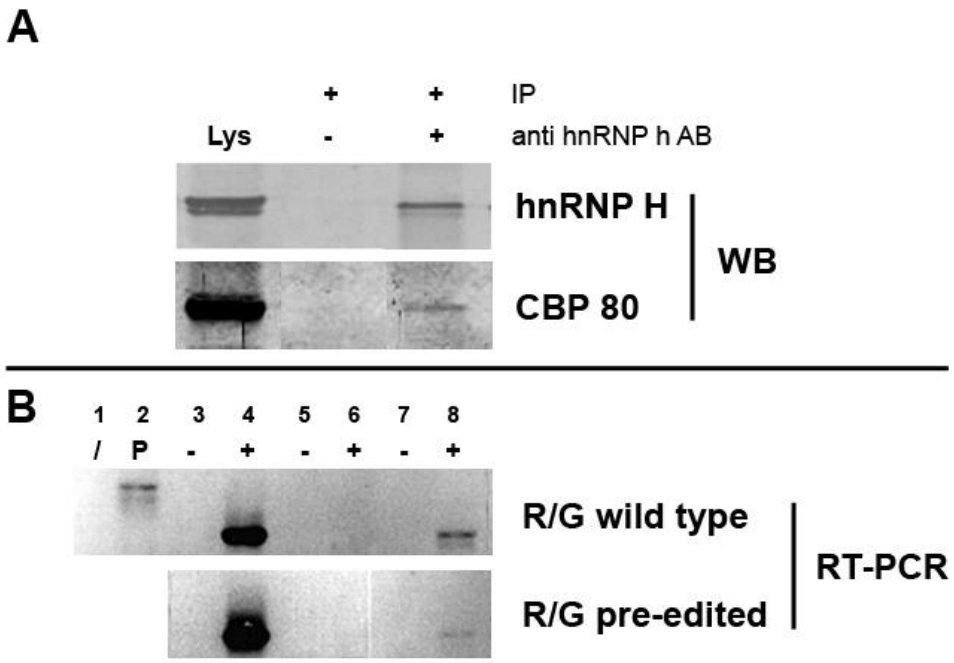

Figure 4.8: Co-immunoprecipitation of R/G site pre-mRNAs with hnRNP H

A: Western Blot for hnRNP H IP detection: hnRNP H was successfully immunoprecipitated (upper panel, lane 3); CBP 80, a known interaction partner of hnRNP H, was used as a positive control for the IP efficiency (lower panel, lane 3); total cell lysate was used as a positive control for WB (lanes 1); uncoupled beads were used as negative control (lanes 2).

B: RT-PCR of co-immunoprecipitated RNAs: the more efficiently spliced wild type construct is found more strongly associated with hnRNP H than the pre-edited construct (spliced variants; compare lanes 8 of upper and lower panels); as positive controls, plasmid DNA (lane 2, indicating the size of unspliced product) and total cell lysate RNA (lane 4, indicating the size of the fully-spliced product) were used as PCR templates (upper panel); "-"lanes represent PCR samples without RT enzyme in the reverse transcription reaction; n.b., all samples are placed exactly under their corresponding IP fractions from part A of this figure.

To our surprise, the association of the wild type R/G site construct with hnRNP H was stronger than the one of the pre-edited construct (Figure 4.8). In both cases, only the correctly spliced product could be found associated with hnRNP H. Then, we went on to see if any major difference could be observed in the association of these pre-mRNAs with hnRNP A1 (Figure 4.9).

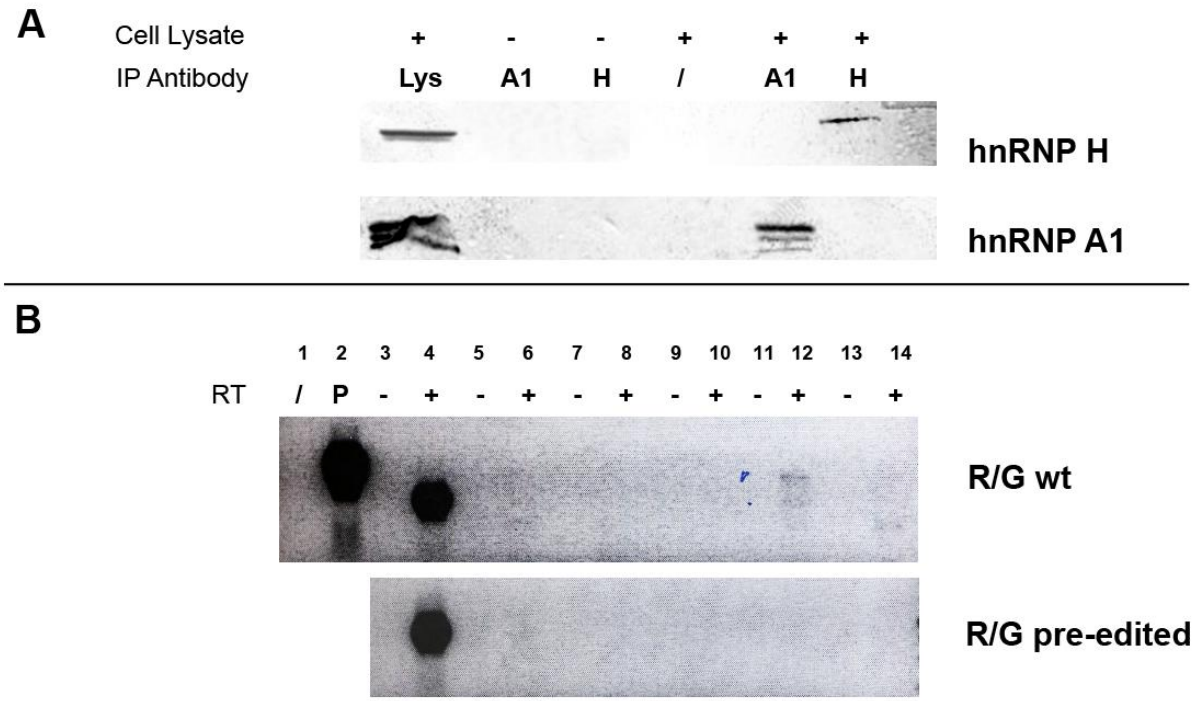

Figure 4.9: Co-immunoprecipitation of R/G site pre-mRNAs with different hnRNPs

A: Western Blot for hnRNP IP detection: two different hnRNPs (hnRNP H, upper panel; hnRNP A1, lower panel) were successfully immunoprecipitated (lanes 5 and 6, respectively); total cell lysate was used as a positive control for WB (lane 1); antibody purity and cross-reactivity was tested for both anti-hnRNP H and hnRNP A1 in lanes 2 and 3, respectively; uncoupled beads were used as negative control (lane 4).

B: RT-PCR of co-immunoprecipitated RNAs: the more efficiently spliced wild type construct is found associated with both hnRNP A1 (predominantly unspliced version; upper panel, lane 12) and hnRNP H (spliced variants; upper panel, lane14); for the pre-edited construct, no such association could be observed (lanes 12 and 14, lower panel); as positive controls, plasmid DNA (lane 2, indicating the size of unspliced product) and total cell lysate RNA (lane 4, indicating the size of the fully-spliced product) were used as PCR templates (upper panel); "-"lanes represent PCR samples without RT enzyme in the reverse transcription reaction; all samples are placed exactly under their corresponding IP fractions from part A of this figure.

Contrary to our original hypothesis, we could not detect any enrichment of pre-edited messages in hnRNP A1- or hnRNP H-coimmunoprecipitated fractions, when compared to the wild type RNAs. In fact, no interaction whatsoever of the aforementioned hnRNPs with the pre-edited mRNAwas observed (Figure 4.9 B, lower panel, lanes 12 and 14). Interesting feature of the experiment was that the unspliced wild type mRNA was enriched in hnRNP A1 fraction (Figure 4.9 B, upper panel, lane 12), while the spliced version was found preferentially associated with hnRNP H (Figure 4.9 B, upper panel, lane 14).

4.2 ADAR1-mediated export of miRNA precursors

The existence of both 110 kDa constitutively expressed (ADAR1-c or ADAR1p110) and the 150 kDa interferon-inducible shuttling isoform (ADAR1-i or ADAR1p150) of ADAR1, synthesised by translation initiation at alternative methionine codons, lead to a proposal that the role of the latter is in hyperediting of viral RNAs. However, upon the emergence of RNA interference and the establishment of its intimate connection with RNA editing, especially the RNA editing of the primary miRNA transcripts, the question of association of ADAR1-i and miRNA precursors during the export process was soon raised. Moreover, full length version of ADAR1 binds siRNAs with an exceptionally strong affinity (Yang et al., 2005).

The idea of ADAR1-i escorting miRNA precursors through the nuclear pores was further encouraged after the identification of Exportin5, a Crm-1 related karyopherin known to interact with dsRBD-containing proteins, as the main export receptor of miRNA precursors (Brownawell and Macara, 2002; Fritz et al., 2009; Lund et al., 2004).

First two microRNAs reported as editing targets were miR-22 and miR-142. Editing of primary miR-22 transcripts occurs in many tissues at relatively low levels. This scarcity of detectable microRNA editing products, however, could also be attributed to their increased susceptibility to degradation (Luciano et al., 2004). Precursors of hematopoietic tissue-specific miR-142, on the other hand, lose their silencing potential upon RNA editing, judged by the decreased Drosha-DGCR8 processivity on the inosine-containing primary miR-142 and by the upregulation of mature miR-142 levels in spleen and thymus of both ADAR1 and ADAR2 *null* mice. Additionally, highly edited primary miR-142 transcripts, in contrast to the unedited ones, were prone to *in vitro* cleavage by exclusively cytoplasmic ribonuclease, Tudor-SN, specific for inosine-containing dsRNA and previously identified as a RISC component (Yang et al., 2006). The residual enigma, how the solely nuclear primary transcript reaches Tudor-SN in the cytoplasm, remained to be clarified. Another open question was whether RNA editing of miRNA precursors was restricted to primary transcripts only, or could pre-miRNAs also be directly affected by this posttranscriptional modification (O'Connell and Keegan, 2006).

This is why I set out to unveil interaction between ADAR1 and established RNA editing targets among miRNA precursors in these two cellular compartments.

4.2.1 ADAR1p150 and miR-22 precursor interact in both nucleus and cytoplasm

The obvious experimental approach to detect RNA-protein interactions, in this case the association of ADAR1 and miR-22 precursor, is co-immunoprecipitation. In our lab, various antibodies against ADAR1 were raised. For our IP experiments, antibodies labelled Sat 309 and Sat 1146 (both raised in rabbits) were coupled to Protein A Sepharose and applied on cell lysates in attempt to pull down ADAR1 miRNA targets. While a peptide from the central region of the protein was used to raise Sat 309, enabling it to recognise both ADAR1p110 and ADAR1p150, N-terminal fusion protein containing the amino-terminus of the full-length ADAR1 served as an antigen in Sat 1146 production, thus allowing the detection of ADAR1p150 only.

Nuclear and cytoplasmic fractions of confluent HeLa cells were prepared by homogenisation. Then, the fraction lysates were made and applied to PAS-coupled antibodies for immunoprecipitation. The precipitated proteins and RNAs were detected by Western Blot and RT-PCR, respectively.

As seen in Figure 4.10 A, Sat 309 antibody is able to recognise and precipitate the corresponding versions of ADAR1. Judging by the RT-PCR experiments done on the immunoprecipitated material in Figure 4.10 B, either both or just one version of ADAR1 interact with miR-22 precursor, a previously identified editing target, in both cytoplasm and the nucleus. The control miRNA precursor, miR-21, on the other hand, could not be detected in the immunoprecipitated fractions by means of RT-PCR, corroborating the assumption that the interaction of miRNA precursors with RNA editing enzyme does not occur randomly, but is limited to the actual editing targets among this class of small RNAs. To control for the separation of nuclei from the cytoplasm, RT-PCR of fully processed (i.e. exclusively cytoplasmic) γ-tubulin B-subunit transcript was employed (Figure 4.10 B, lower panel).

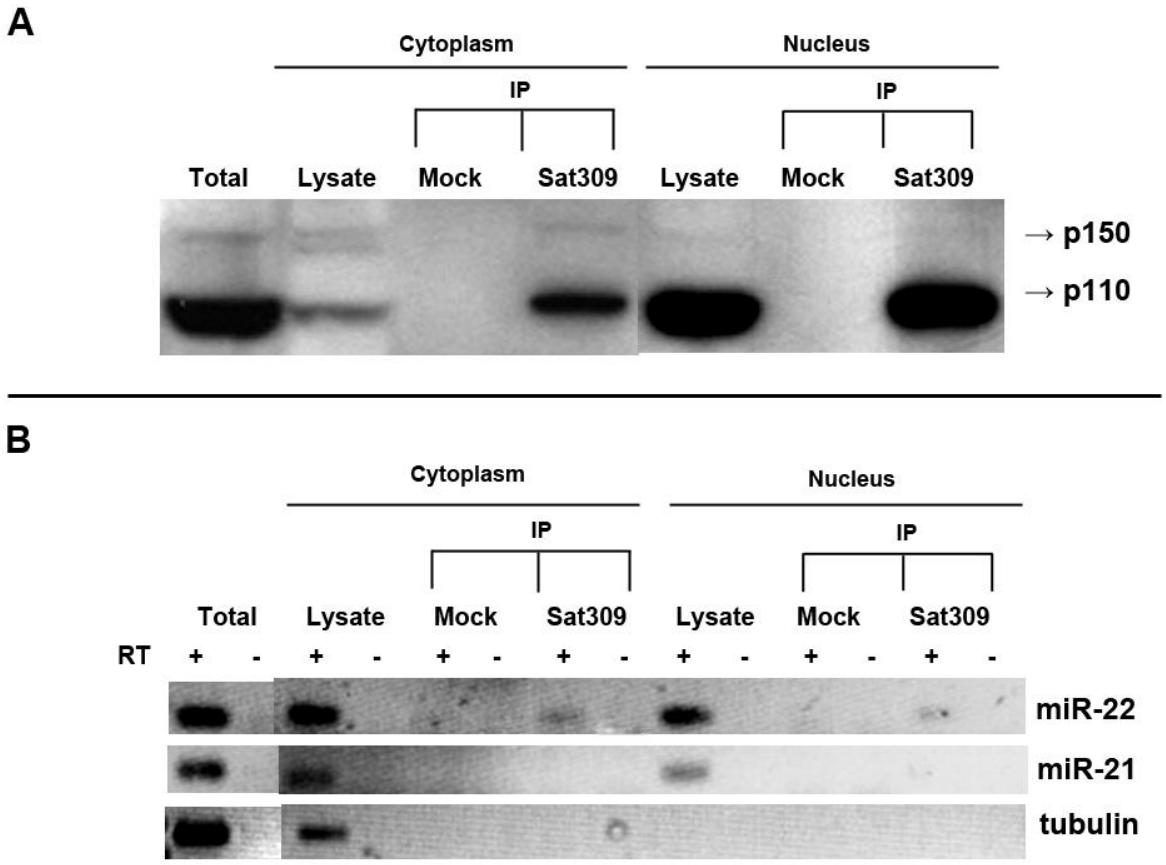

Figure 4.10: Co-immunoprecipitation of ADAR1p110 and miR-22 precursor

A: Western Blot used for ADAR1 IP detection: nuclear predominance of short (p110) and cytoplasmic predominance of long (p150) ADAR1 can be clearly seen (see Cytoplasm and Nucleus Lysate lanes); total cell lysate was used as a control (Total); adequate immunoprecipitation conditions were met, as seen in Sat309 lanes; uncoupled beads were used as negative control (Mock); for Western Blot detection, Sat12 antibody which detects both ADAR1 versions was used.

B: RT-PCR of co-immunoprecipitated RNAs: miR-22, but not miR-21 precursor can be co-immunoprecipitated with ADAR1 from both cytoplasmic and nuclear lysate ("+" lanes Sat309 for both cytoplasmic and nuclear fractions; "-" lanes indicate negative controls for reverse transcription, samples where no enzyme was added to the RT reaction); γ-tubulin B-subunit mRNA RT-PCR served as a control of separation of nuclei from the cytoplasm (lower panel); * lanes correspond to the ones denoted in A

Next, we wanted to see if both ADAR1 versions interact with miR-22 precursors or if it is restricted to ADARp150. This is why we used ADAR1p150-specific antibody Sat 1146 for the next round of co-immunoprecipitation experiments (Figure 4.11).

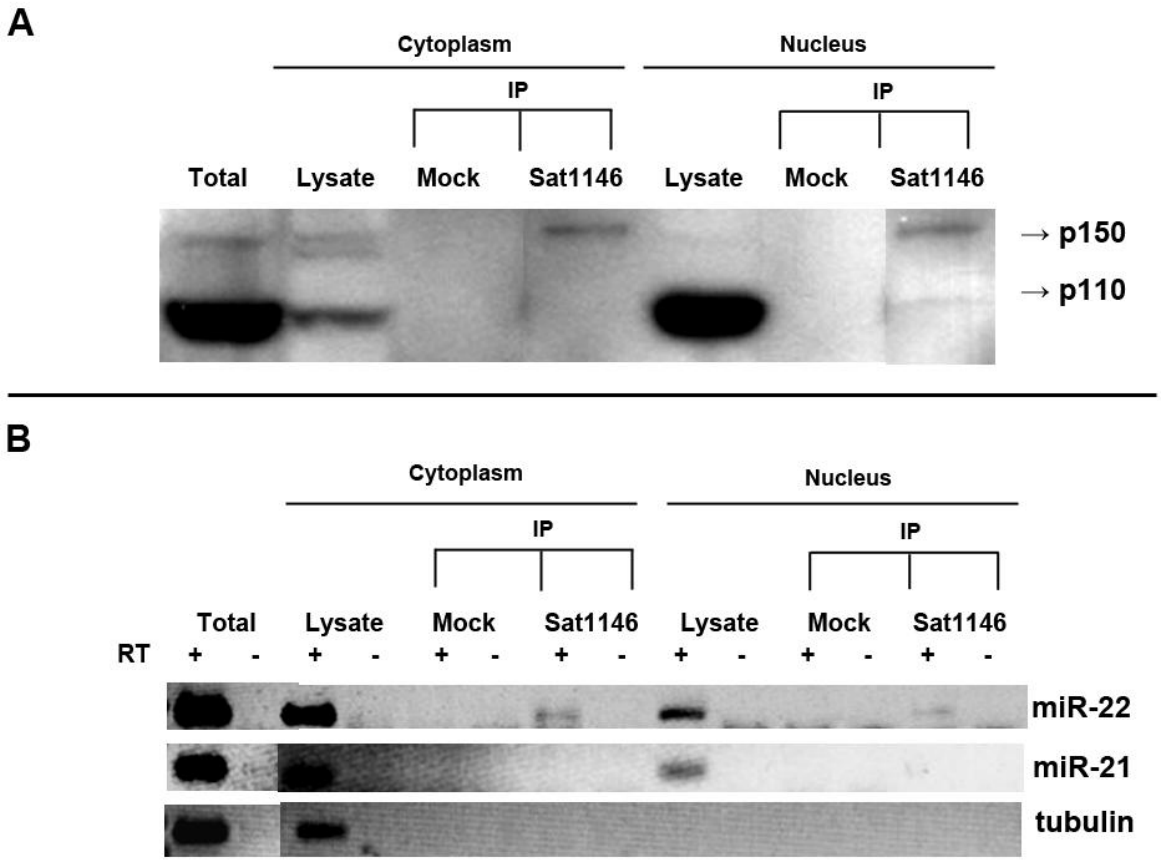

Figure 4.11: Co-immunoprecipitation of ADAR1p150 and miR-22 precursor

A: Western Blot used for ADAR1 IP detection: total cell lysate was used as a control (Total); adequate immunoprecipitation conditions were met, as judged by Sat1146 lanes; uncoupled beads were used as a negative control (Mock); for Western Blot detection, Sat12 antibody which detects both ADAR1 versions was used.

B: RT-PCR of co-immunoprecipitated RNAs: miR-22, but not miR-21 precursor can be co-immunoprecipitated with ADAR1p150 from both cytoplasmic and nuclear lysate ("+" lanes Sat1146 for both cytoplasmic and nuclear fractions; "-" lanes indicate negative controls for reverse transcription, samples where no enzyme was added to the RT reaction); γ-tubulin B-subunit mRNA RT-PCR served as a control of separation of nuclei from the cytoplasm (lower panel); * lanes correspond to the ones denoted in A.

Indeed, the RT-PCR signals obtained for the miR-22 precursor co-immunoprecipitated with ADAR1p150 showed similar intensity to the ones derived from the material co-immunoprecipitated with both ADAR1 isoforms (Figure 4.11, panel B).

This is why we conclude that the fraction of co-IP-ed miR-22 precursor was comparable using either antibody (no enrichment when using Sat 309), indicating that the interaction is probably restricted to the full-length, shuttling ADAR1 isoform (ADAR1p150).

The co-immunoprecipitation experiment clearly shows the specific association of ADAR1p150 with miR-22 precursor in both nucleus and the cytoplasm, supporting our original hypothesis according to which miRNA precursor editing targets serve as RNA bridges between dsRBD protein ADAR1p150 and Exportin5 during nuclear export.

The next step to further corroborate this working model was to check for the differences in editing levels of different miR-22 precursors and mature form in lysates versus ADAR1p150 immunoprecipitates from nuclear and cytoplasmic fractions, expecting to find a gradual increase of inosine content towards the mature miRNA forms, in the cytoplasm. This approach, however, had to be abandoned because of the technical difficulties caused by contamination in RT-PCR.

Similar experimental concept was planned for another pronounced editing target among miRNA precursors, namely miR-142, expressed mainly in hematopoietic tissues. Unfortunately, due to time restrictions this project was put on hold.

4.3 SINE RNA interacts with HYL1, key member of RNAi pathway in plants

A collaboration with the lab of Jean-Marc Deragon at the University of Perpignan, France was initiated to investigate the role of tRNA-derived SINE RNAs in the RNAi pathway in plants.

Transgenic *A. thaliana* lines overexpressing SB1 SINE RNA exhibit developmental defects similar to the ones seen in some RNAi mutants. In *Arabidopsis*, SB1 transcription is associated with delayed growth and flowering time, reduction of leaf and root size, etc. Several of these defects resemble those observed in *hyl1* and *drb4* mutants, which are impaired in the two dsRBPs required for miRNA and trans-acting small interfering RNA (tasiRNA) pathways, respectively, suggesting that SB1 could interact with RNA-binding proteins of the miRNA or tasiRNA pathways (Pouch-Pelissier et al., 2008).

Moreover, SB1 RNA does not conserve the ancestral tRNA folding pattern, but instead adopts a structure consisting of three stem-loops with bulges and mismatches (Sun et al., 2007). This, in turn, raises the possibility of interaction with various dsRBPs (double-stranded RNA binding proteins) given that the recognition of dsRNA by dsRBPs generally does not involve sequence specificity and several structured RNAs forming stem-loops with bulges or mismatches were shown to bind efficiently to dsRBPs, many of which are involved directly in RNAi (Stefl et al., 2005). In *Arabidopsis*, these proteins include the four Dicer-like proteins (DCL1 to 4), the five dsRNA-binding proteins (HYL1 and DRB2 to 5) and the HUA Enhancer 1 (HEN1) protein (Hiraguri et al., 2005).

Because the production of SINE RNA induces growth anomalies reminiscent of those in *hyl1* and *drb4* null mutants, we hypothesised that if SB1 RNA mimicked the structure of natural mi/tasiRNA substrates, it could interact with and titrate proteins involved in the biogenesis of these small RNAs So, in order to explain the observed similarity between SB1 expressing lines and RNAi mutants, the affinity of SB1 SINE RNA for HYL1 and DRB4 was tested.

4.3.1 SINE RNA interacts with a subset of double-stranded RNA binding proteins

HYL1 is part of the DCL1 complex and is involved in processing miRNA primary transcripts (pri-miRNAs) and short precursors (pre-miRNAs). In gel retardation experiments, SB1 SINE RNA, but not DNA or single-stranded RNA fragments of a similar size, associated with a recombinant GST-HYL1 fusion protein (Pouch-Pelissier et al., 2008). Although a perfect RNA duplex also could bind HYL1, this association was less efficient compared to SB1 RNA, suggesting that HYL1 prefers dsRNA substrates containing bulges and/or distal loops. RNase V1 protection assay helped us define more precisely the SB1 RNA binding sites of HYL1 (Figure 4.12): HYL1 binds mainly to the first and longest SB1 stem-loop, adopting a fold similar to pre-miRNAs. Indeed, the protected region includes an RNA duplex containing mismatches and extends into the single-stranded terminal loop region. A structurally similar, although weaker HYL1 binding site also is present on the second stem-loop of SB1 RNA. We also tested the capacity of DRB4, a dsRBP involved in the production of tasiRNAs (Mallory and Vaucheret, 2006) to bind SB1 RNA. In this case, the GST-DRB4 fusion protein did not bind significantly to SINE RNA in our *in vitro* assay, although it did efficiently bind to a perfect RNA duplex, which likely resembles the structure of tasiRNA templates. This result suggests that SINE RNAs interact with dsRBPs specifically adapted to bind imperfect double-stranded RNA rather than those that bind perfect RNA duplexes.

Recently, dsRBDs from two *Xenopus laevis* proteins, xlADAR1 and xlRPBA, were shown to bind efficiently to short stem-loop RNA structures containing bulges and mismatches (Hallegger et al., 2006), consistent with the observation that ADARs can modify miRNA precursors *in vivo* (Kawahara et al., 2007a). Because of structure similarity, it is possible that many SINE RNAs interact with dsRBPs. We performed binding experiments with SB1 RNA and the second double-stranded RNA binding domain of xlADAR1 (called Dr2) and mapped the RNA binding sites (Figure 4.12). We observed that Dr2 bound SB1 RNA in the same region as HYL1, suggesting that SINE RNAs have the potential to interact with a subset of dsRBPs across eukaryotic species, including the ones involved in miRNA production. DRB4 had no impact on RNase V1 cleavage pattern, confirming its inability to bind SB1.

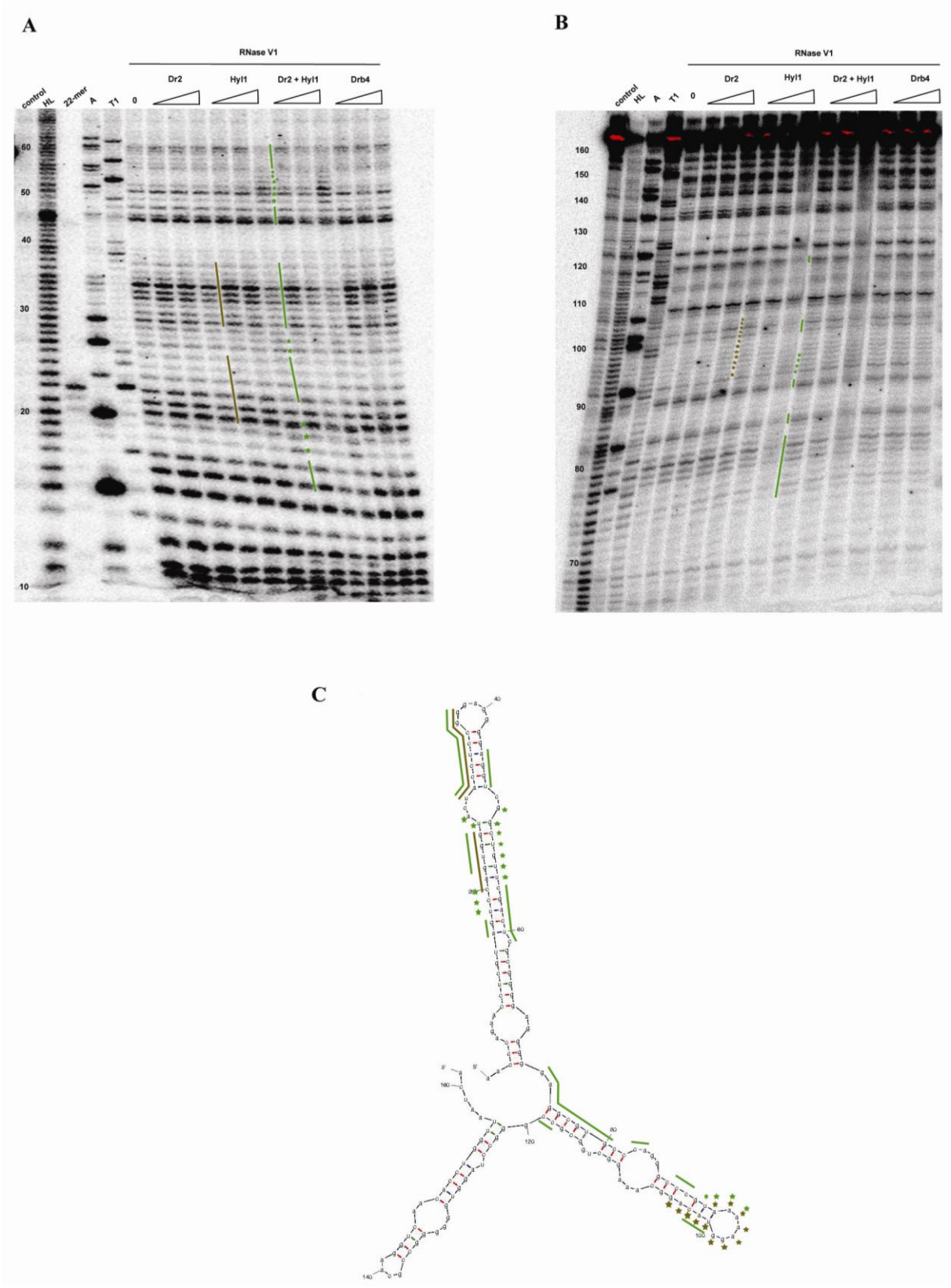
A
RNase V1
control
HL
22-mer
A
T1
0
Dr2
Hyl1
Dr2 + Hyl1
Drb4
60
50
40
30
20
10
B
RNase V1
control
HL
A
T1
0
Dr2
Hyl1
Dr2 + Hyl1
Drb4
160
150
140
130
120
110
100
90
80
70
C

Figure 4.12

Figure 4.12: Protection from RNase V1 digestions of SB1 SINE RNA by two different dsRBPs

Prior to RNase V1 digestion, *in vitro* transcript of SB1 was subjected to protection by increasing concentrations of expressed dsRBPs: Dr2 (the second dsRBD of *Xenopus laevis* ADAR1), HYL1, combination of Dr2 and HYL1, or DRB4. Regions protected by Dr2 and HYL1 are marked alongside short run gel (**A**), long run gel (**B**) and predicted folding pattern (**C**) by brown and green bars, respectively. Three independent experiments gave similar results as the one presented. Nucleotides marked with asterisks seem to adopt more prominent helical structure upon protein binding and, therefore, become more prone to RNase V1 cleavage. DRB4 is showing no effect on RNase V1 cleavage, confirming its low *in vitro* binding affinity to SB1 RNA. Dr2 and HYL1, in this case, bind to and protect similar regions of RNA. HYL1 is, however, showing stronger binding affinity than Dr2. (HL) represents a partially hydrolyzed RNA ladder. Denaturating RNase A and T1 digests give the position of pyrimidine and G residues respectively. The control lane shows untreated RNA samples and the (0) lane represent RNase V1 digestion without recombinant proteins added. A labeled 23-mer oligoribonucleotide was also loaded on the gel to help in band size determination.

4.3.2 HYL1 induces conformational changes upon binding to RNA

The binding of HYL1 appears to increase the RNase V1 sensitivity of certain regions of the SB1 RNA (indicated by asterisks on Figure 4.12 A). The RNase V1 activity is sensitive to RNA conformation and, although sensitivity does not always imply hydrogen bonding of the bases in a canonical double stranded helix, it does require a structured, helix-like conformation. As such, the increased RNase V1 sensitivity following HYL1 binding suggests that HYL1 is able to force some single-stranded RNA regions to adopt a more structured helix-like conformation, possibly by promoting non Watson-Crick base pairing.

To test a chaperon-like activity for HYL1 and to explore its generality, we performed binding experiments using the SELEX clone 11Dr2(7), a short imperfect double-stranded RNA known to bind the Dr2 motif. Following RNase V1 digestion, we confirmed the binding of Dr2 to 11Dr2(7) (Figure 4.13). We observed that HYL1 is able to bind strongly to 11Dr2(7) and generate regions protected from RNase V1 activity (represented by green lines on Figure 4.13 B) and regions with increased sensitivity to RNase V1 (represented by asterisks on Figure 4.13 A and B). Again DRB4 was unable to bind 11Dr2(7) and no synergetic effect was observed when HYL1 and Dr2 were used together. Similar results were observed when the SB2 *Arabidopsis* SINE RNA was used as a substrate (see Figure 4.14). Our results suggest that, upon binding RNA, HYL1 has the general capacity to force single-stranded regions to adopt a more organized, helix-like configuration.

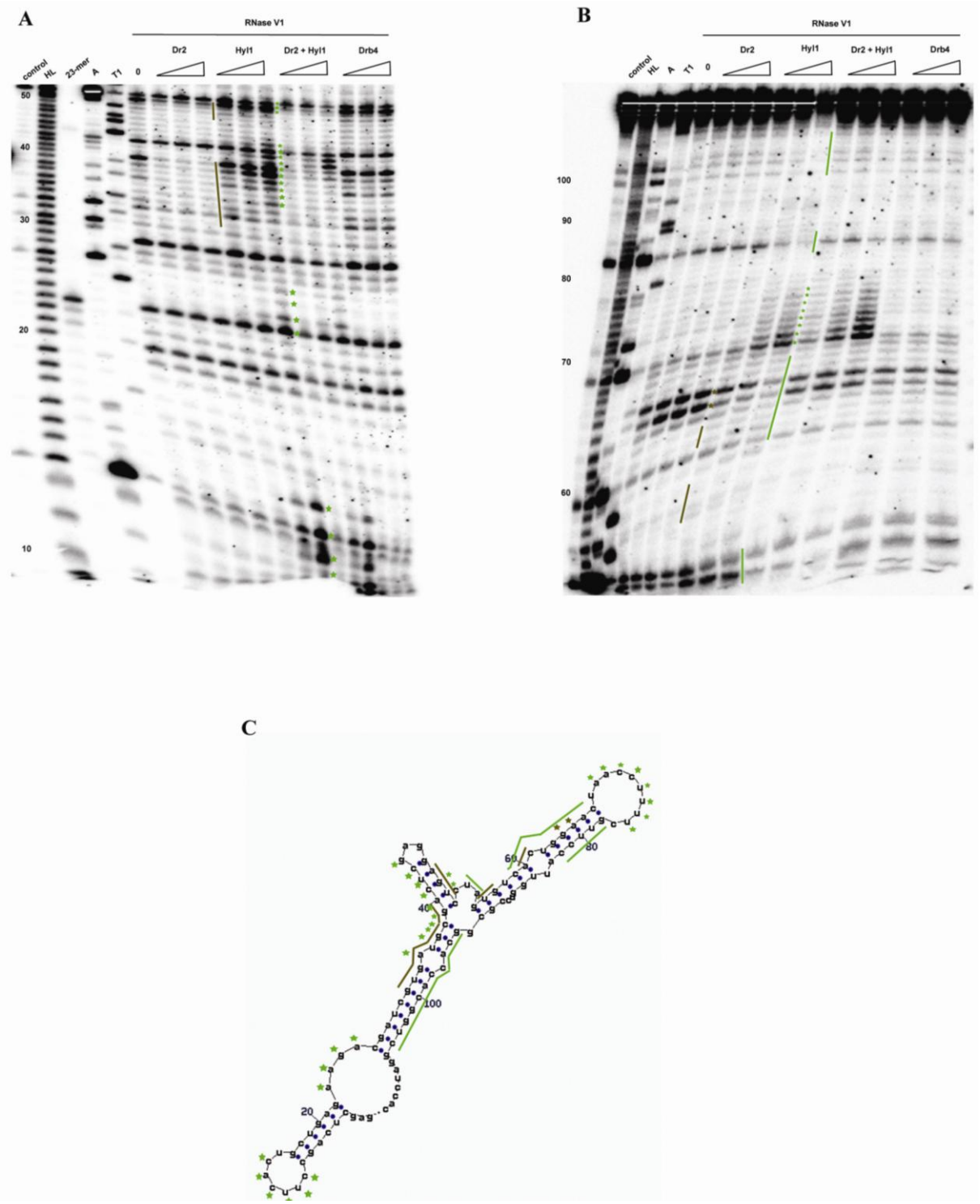
A
RNase V1
Dr2
Hyl1
Dr2 + Hyl1
Drb4
control
HL
23-mer
A
T1
0
50
40
30
20
10
B
RNase V1
Dr2
Hyl1
Dr2 + Hyl1
Drb4
control
HL
A
T1
0
100
90
80
70
60
C
20
40
60
80
100

Figure 4.13

Figure 4.13: Protection from RNase V1 digestion of SELEX clone 11Dr2(7) by different dsRBPs.

Prior to RNase V1 digestion, *in vitro* transcript of SELEX clone 11Dr2(7) was subjected to protection by increasing concentrations of expressed dsRBPs: Dr2, HYL1, combination of Dr2 and HYL1, or DRB4. Regions protected by Dr2 and HYL1 are marked alongside short run gel (**A**), long run gel (**B**) and predicted folding pattern (**C**) by brown and green bars, respectively. Three independent experiments gave similar results as the one presented. Nucleotides marked with asterisks seem to adopt more prominent helical structure upon protein binding and, therefore, become more prone to RNase V1 cleavage. DRB4 is showing no effect on RNase V1 cleavage of given RNA, confirming its low binding affinity for imperfect RNA duplexes. (HL) represents a partially hydrolyzed RNA ladder. Denaturating RNase A and T1 digests give the position of pyrimidine and G residues respectively. The control lane shows untreated RNA samples and the (0) lane represent RNase V1 digestion without recombinant proteins added. A labeled 23-mer oligoribonucleotide was also loaded on the gel to help in band size determination.

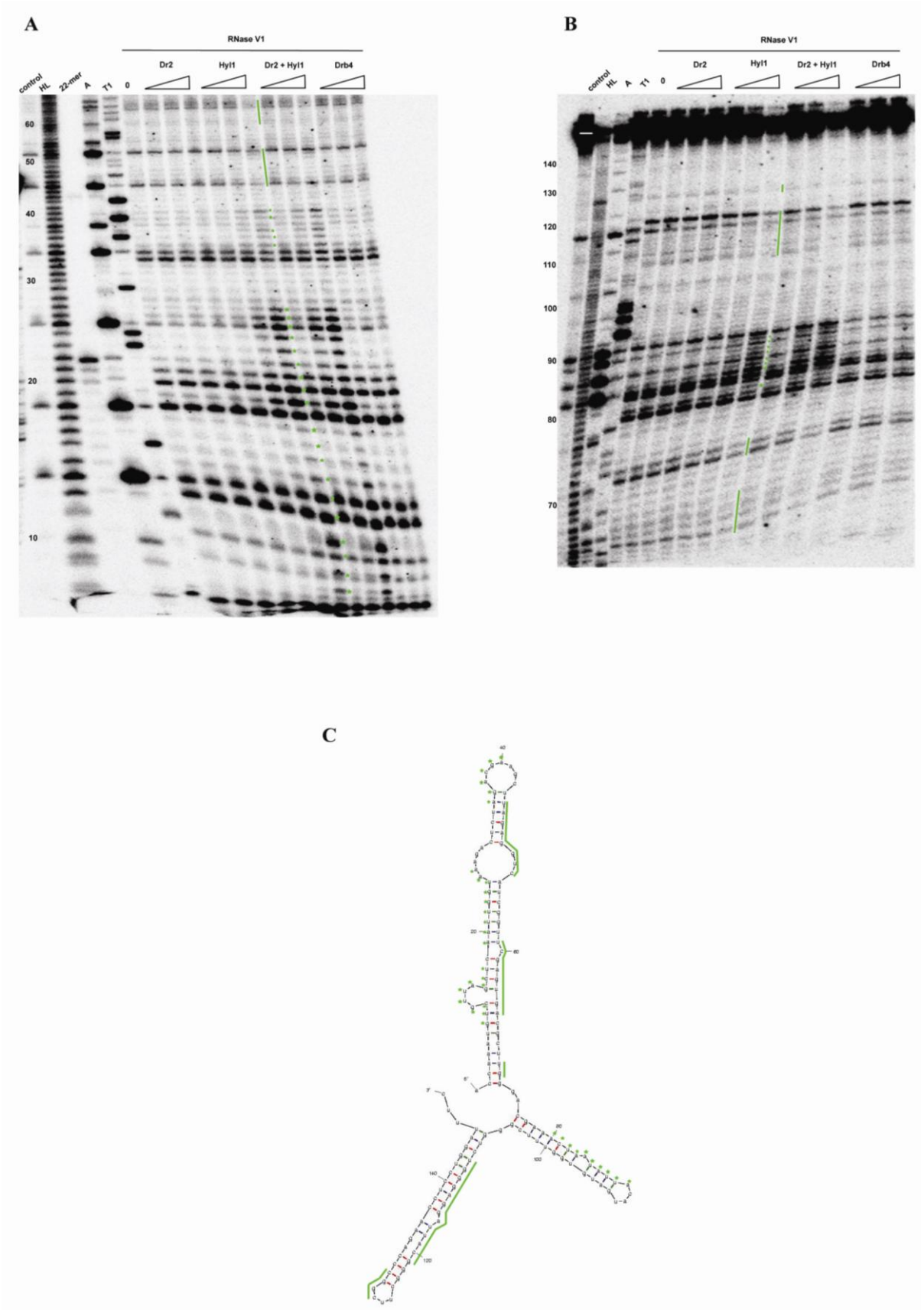
A
RNase V1
control
HL
22-mer
A
T1
0
Dr2
Hyl1
Dr2 + Hyl1
Drb4
60
50
40
30
20
10
B
RNase V1
control
HL
A
T1
0
Dr2
Hyl1
Dr2 + Hyl1
Drb4
140
130
120
110
100
90
80
70
C

Figure 4.14

Figure 4.14: Protection from RNase V1 digestion of SINE SB2 RNA by different dsRBPs

Prior to RNase V1 digestion, *in vitro* transcript of SB2 was subjected to protection by increasing concentrations of following dsRBPs: Dr2, HYL1, combination of Dr2 and HYL1, or DRB4. Regions protected by HYL1 are marked alongside short run gel (A), long run gel (B) and predicted folding pattern (C) by green bars. Three independent experiments gave similar results as the one presented. Nucleotides marked with asterisks seem to adopt more prominent helical structure upon protein binding and, therefore, become more prone to RNase V1 cleavage. (HL) represents a partially hydrolyzed RNA ladder. Denaturating RNase A and T1 digests give the position of pyrimidine and G residues respectively. The control lane shows untreated RNA samples and the (0) lane represent RNase V1 digestion without recombinant proteins added. A labeled 23-mer oligoribonucleotide was also loaded on the gel to help in band size determination. The binding of HYL1 to SB2 is weaker compared to SB1 or 11Dr2(7). In this case, both Dr2 and DRB4 are showing no effect on RNase V1 cleavage, suggesting their low *in vitro* binding affinity to SB2 RNA.

5 DISCUSSION

5.1 Compensatory U1 snRNA partially restores the splicing efficiency of pre-edited R/G site construct

ADAR family of enzymes catalyses hydrolytic deamination of adenosines to inosines within double-stranded RNAs. In mRNAs, this reaction can take place in exons, introns or untranslated regions (UTRs), leading to codon exchange, splice site introduction or stop codon abolishment.

A → I RNA editing in exons can result in codon exchange and subsequent amino acid substitution. Codon exchange as a direct consequence of RNA editing in mammals has first been reported for the AMPA receptor subunit GluR-B (Higuchi et al., 1993). Adenosine deamination occurs at multiple sites in this particular transcript: in exon 11 (Q/R site), in intron 11 (editing hotspots 1 and 2) and, ultimately, in exon 13, just two nucleotides upstream of the intron 13 5´splice site (R/G site). Editing at the R/G site causes a reduction in intron 13 splicing efficiency (Schoft et al., 2007). It is postulated that this transient arrest is necessary to ensure that proper alternative splicing decisions for the downstream, mutually exclusive exons 14 and 15 are made: underediting at the R/G site, apparently, could lead to hasty activation of the branch point sites (BPSs) in introns 13, 14 and 15, resulting in either inclusion of both exon 14 and 15, or in skipping them entirely. This, in turn, would have the emergence of premature stop codons and synthesis of non-functional receptors as a consequence. A similar effect of underediting on alternative splicing was perceived for the mammalian serotonin 5-HT_{2C} receptor (Maas et al., 2001) and for the number of edited mRNAs in *Drosophila melanogaster* (Agrawal and Stormo, 2005). On the other hand, there is evidence that RNA editing stimulates developmentally regulated alternative splicing events (Gallo et al., 2002).

The explicit step in the complex pre-mRNA splicing network that is affected by RNA editing, however, was not yet determined. At the R/G site, deamination occurs at a conserved -2 position (with respect to the 5´splice site). This adenosine is involved in the base-pairing interaction with U1 snRNA in one of the very first steps of the spliceosome assembly, which lead us to speculate that mRNA:U1snRNA interaction gets profoundly destabilised once inosine comes into play. This is why the imperfect 5´splice site hypothesis was tested by employing compensatory U1 snRNA mutants to investigate the influence of inosine in this critical position on U1 snRNA base-pairing potential.

On the other hand, the sequence surrounding the edited R/G site bears strong resemblance to the hnRNP A1-binding or to the G-rich hnRNP H-binding splice silencer consensus sequences. It, therefore, seems also plausible that RNA editing creates a binding site for one of the hnRNPs which mediate splicing repression, explaining the observed effect on splicing. In this light, the question of association of these two hnRNPs with target mRNA in different editing states was addressed.

5.1.1 U1 snRNA interaction with 5´ splice site is disrupted upon RNA editing at R/G site

In this work, we focused on identifying the macromolecular interaction within the RNA splicing machinery that gets impaired once the co-transcriptional adenosine-to-inosine RNA editing occurs in close proximity to 5´ splice site, resulting in reduced splicing efficiency, just as in the case of the R/G site editing in exon 13 of the GluR-B pre-mRNA.

Using the *in vivo* splicing assay established in our lab, we are able to show that the introduction of the compensatory mutation into U1 snRNA in the nucleotide position forming a base-pair with the editied nucleotide in the pre-mRNA could increase the splicing competence of the affected RNA to some extent. The lack of the full restoration of the wild type splicing efficiency in this experimental system could be attributed to either lower transfection rates for the U1 snRNA-bearing constructs or the less favourable incorporation of the mutated U1 snRNA into U1 snRNP compared to the wild type small nuclear RNA, or both. It is also plausible that yet another, thus far unidentified, step in pre-mRNA maturation is affected by RNA editing of the 5´ splice site adjacent sequence. In any case, other experimental approaches that included overexpression of exogenous small nuclear RNAs like U1 or U2 snRNA also resulted in expression levels similar to the ones reached in our experiments (Pan and Prives, 1988; Zhuang and Weiner, 1986; Zhuang and Weiner, 1989).

The role of the base-pairing interaction between the 5´ end of U1 snRNA and the 5´ splice site in pre-mRNA is indispensable for the 5´ splice site recognition by the splicing machinery (Zhuang and Weiner, 1986). However, before the recognition step may take place, a complex network of protein-protein and protein-RNA interactions has to be formed, most notably within the U1 snRNP and its interaction partners (Kohtz et al., 1994). U1 snRNA itself adopts extensive secondary structure consisting of four stem-loops and two relatively short single-stranded regions (Krol et

al., 1990). The first single-stranded stretch is found at the 5´ end and is responsible for the interaction with the 5´ splice site, while the other one represents the Sm-binding site between stem-loops 3 and 4, around which seven Sm proteins assemble (Urlaub et al., 2001). Other protein components of U1 snRNP, namely U1-70K and U1-A use their RNA-binding domains to interact with stem-loops 1 and 2 of U1 snRNA, respectively (Oubridge et al., 1994; Patton and Pederson, 1988). In contrast, U1-C protein is not able to bind free U1 snRNA, but is only capable of achieving this interaction once it establishes protein-protein contacts with U1-70K (Nelissen et al., 1994).

Recently, the crystal structure of human U1 snRNP was solved at 5.5 Å resolution (Pomeranz Krummel et al., 2009). The structure of U1 snRNP implies an involvement of U1-C in stabilising interaction between the 5´ end of U1 snRNA and the 5´ splice site in pre-mRNA, judged by the data showing that the loop region and helix A of the U1-C zinc finger domain bind to the RNA duplex mimicking the base-paired U1 snRNA::5´ splice site dsRNA. Moreover, U1-C is crucial for the formation of the early spliceosomal complex, as the double mutant of U1-C fails to promote it (Heinrichs et al., 1990; Will et al., 1996). The binding occurs in the minor groove of RNA duplex where C8 and A7 nucleotides of U1 snRNA base-pair with conserved GU dinucleotide of the 5´ splice site, being highly likely that the adjacent base-pairs, among them the one formed by -2 nucleotide of the 5´ splice site (R/G site) and nucleotide 10 of U1 snRNA, also play a role in this protein-dsRNA association. The authors suggest that the conserved residues in the loop and helix A region of U1-C could serve to probe the hydrogen bonding of the base pairs in the minor groove of RNA duplex and, subsequently, allowing U1-C binding only to a correctly paired 5´ splice site and U1 snRNA would promote the following step in the splicing network.

These propositions would also go along the same lines with the results we obtained in our studies and could explain the observations we came across. In our case, the inosine/guanosine instead of adenosine at the -2 position of the 5´ splice site would not be able to form a stable base pair with the U10 of U1 snRNA. In addition, U1-C would sense the unfavourable thermodynamic conditions, evading binding and failing to promote the formation of the complex required for the next step in RNA splicing. This would result in the decrease of the splicing efficiency, observed by us as reduction in green fluorescence intensities of our construct. On the other hand, introduction of the compensatory mutation in U1 snRNA (C10 instead of U10)

and the incorporation of the mutant into U1 snRNP would allow for the formation of stable base pair with the inosine/guanosine at the -2 position of the 5´ splice site. U1-C could form a strong association with RNA duplex in this scenario, promoting further steps in the spliceosome assembly. This effect was displayed as the partial restoration of the original splicing efficiency in our experiments.

5.1.2 Role for hnRNPs in regulation of splicing through RNA editing

Alternative splicing of pre-mRNAs is a widespread phenomenon that contributes considerably to proteome variety in higher organisms. In mammals, for example, majority of genes undergo this process, often on a tissue-specific level, through alternative exclusion or inclusion of either introns (intron retention) or exons (exon skipping) during mRNA processing (Wang et al., 2006).

On the molecular level, splicing enhancer signatures in mRNAs are generally bound by the serine/arginine (SR) rich proteins , which then promote the recognition of adjacent splice sites (Wang et al., 2004b). Splice silencer sequences, on the other hand, are most commonly recognised by one of the members of the hnRNP protein family, including hnRNP I/PTB (polypyrimidine tract-binding protein), hnRNP A1 or hnRNP H (Chen et al., 1999; Grabowski, 2004). Each of these protein factors has its own subset of the mRNA sequence motifs it associates with preferentially.

In the case of R/G site in GluR-B pre-mRNA, proper alternative splicing decisions are stimulated indirectly only upon RNA editing, which first leads to the splicing reduction and transient intron 13 retention, followed by correct processing of downstream introns (Schoft et al., 2007). The resulting region (UUAIGgugg with intron sequence in lower case) exhibits strong sequence homology to an hnRNP A1-binding splice silencer consensus (GGCAGGGUGG). Moreover, the introduction of guanosine-mimicking inosine into this sequence raises the affinity of another splice silencing regulator, G-rich binding hnRNP H, for this region.

Here, we tried to screen for the differences in pre-edited pre-mRNA binding to two hnRNPs involved in splicing silencing, in comparison to wild type pre-mRNA. Although we expected to see stronger association of the pre-edited message with the previously identified splicing silencers hnRNP H and hnRNP A1, this particular mRNA seemed to associate altogether more weakly or transiently than its wild type counterpart in our co-IP experiments, precluding us from reaching any conclusions on the role of hnRNPs in splicing regulation upon RNA editing.

5.2 ADAR1p150 remains associated with edited miRNA precursor in the cytoplasm

The biogenesis of microRNAs through processing activities in both nucleus and the cytoplasm requires a reliable nuclear export pathway where pre-miRNAs, 60-70 nucleotide long stem-loop products of nuclear Drosha-DGCR8 activity on primary miRNA transcripts (Han et al., 2006; Lee et al., 2003), are delivered to the cytoplasmic Dicer-TRBP complex for the final processing step in which functional mature miRNAs are synthesised, 21-23 nucleotides long single-stranded RNAs capable of silencing their respective messenger RNA targets within the RISC complex. The karyopherin identified as an export receptor for pre-miRNAs was Exportin-5 (Lund et al., 2004). This export receptor was shown to mediate RNA-dependent export of a number of dsRBD proteins. Remarkably, dsRBDs of cargo proteins are in most cases indispensable for the interaction with Exportin-5. In addition, this interaction is thought to be mediated, or at least strengthened, through an RNA bridge (Fritz et al., 2009).

ADAR1 as dsRBP and microRNA precursors have both been shown to interact with Exportin5. Since both nuclear primary microRNAs and cytoplasmic pre-microRNAs were reported as editing targets, we were interested to see if ADAR1p150, being a shuttling protein, escorts these substrates to the cytoplasm.

At the time when this study was concieved, validated editing substrates among miRNA precursors were miR-22 and miR-142, both on the primary transcript level. Nuclear pri-miR-142, though, was shown to be sensitive to cytoplasmic Tudor-SN cleavage, further corroborating our co-export hypothesis (Luciano et al., 2004; Yang et al., 2006).

In this study we present evidence that ADAR1p150 and miRNA-22 precursor associate in the nucleus and continue to interact in the cytoplasm. It provides a strong indication that this is not merely a short-lived complex, but a dynamic macromolecular assembly capable of trafficking through the nuclear pores. Also, the absence of miR-22 precursor enrichment in fractions immunoprecipitated with antibody recognising both full-length and, predominantly nuclear, short ADAR1 isoform suggest that this miR-22 precursor interaction is restricted to ADAR1p150, rationalising its expression and/or shuttling behaviour in non-infected tissues and cells.

miR-142 as the other RNA editing substrate in the class of miRNA precursors was surprisingly reported as ADAR1p110 and ADAR2 target, but not as ADAR1p150 target *in vitro*. In this case, adenosine deamination of nuclear pri-miR-142 at eleven sites leads to conversion of stable A=U Watson-Crick base-pairs into more flexible I-U wobble structures, causing profound alterations in the structure and stability of this dsRNA. This change results in complete blockage of the Drosha cleavage step and reduction in mature miR-142 levels. Out of eleven modified adenosines in this pri-miRNA, the sites identified as the ones leading to inhibition of Drosha-DGCR8 cleavage were positions +4 and +5 (with respect to the mature miR-142-5p sequence), located near the actual Drosha cleavage site. Additionally, this modification provokes an increased sensitivity to Tudor-SN-mediated degradation (Yang et al., 2006). However, this ribonuclease specific for hyper-edited dsRNAs is confined to cytoplasm (Scadden, 2005), again raising the issue of how exactly the edited pri-miR-142 gets out of the nucleus, since RanGTP-Exportin5-dependent pre-miRNA export mechanism applies only to properly processed molecules. On the other hand, the observation that even miRNA precursors longer than properly processed ones (85 nt long miR-22 in this case) can get exported from *Xenopus laevis* nuclei, possibly through an Exportin-5-independent pathway, is compatible with the results we obtained for the miR-22 precursor (Lund et al., 2004).

The next question we wanted to address in this study, but due to time restrictions failed to, was the level of miRNA editing throughout the entire miRNA biogenesis pathway, from the RNA Pol-II-mediated transcription in the nucleus to the mature miRNA delivery to the RISC complex in the cytoplasmic P-bodies. Since we found ADAR1p150 associated with miR-22 precursor also in the cytoplasm, it would be rational to expect the increased editing levels in this cellular compartment, when compared to the nucleus. The experimental approach for this type of analysis would include immunoprecipitation of ADAR1-associated miRNAs and RT-PCR of miRNAs in different maturation stages, followed by cloning and sequencing. A to I changes would then be detected as Gs in the examined sequence. Indeed, in one study a rising inosine content in two intermediates of miRNA formation was reported for a novel miRNA editing target, miR-151, in several mouse tissues. Editing at Dicer cleavage site-adjacent positions -1 and, especially +3 (with respect to the mature miR-151-3p sequence) has proven to act repressive on the Dicer-TRBP processing of pre-miR-151, leading to detection of exclusively unedited mature miR-151

(Kawahara et al., 2007a). On the structural level, the plausible explanation for this inhibition would be the terminal loop extension as the consequence of the stem instability caused by the presence of wobble I-U basepairs, precluding the accurate stem length recognition by Dicer. While editing at +3 position in pri-miR-151 fluctuated between 30-70%, all the detectable pre-miR-151 species (100%) had inosine at this site. *In vitro* experiments have also shown that pre-miR-151 alone gets highly edited by either ADAR1p150 or ADAR1p110, but not ADAR2, consistent with our detection of cytoplasmic ADAR1p150::pre-miR-22 association. This miRNA modification is thought to have a regulatory role in controlling the number of active mature miRNAs through the maturation repression, just like in the case of miR-142.

The most dramatic effect of miRNA editing was the redirection of mature miR-376 to a completely different set of mRNA targets. Here, one highly edited site is situated exactly in the middle of the seed region, responsible for the first steps in recognition and hybridisation of mature miRNAs to their mRNA targets within the RISC complex (Kawahara et al., 2007b).

All these new developments are in line with our preliminary data and encourage further effort in elucidating fate of ADAR1-associated miRNA precursors from the nucleus out to the cytoplasm. Unfortunately, all further efforts to get clean RT-PCR reactions without any genomic DNA contamination for the cDNA sequencing failed for what we believe to be two reasons. Firstly, working with human genes or transcripts enormously increases the contamination risk in such sensitive applications like RT-PCR. Secondly, microRNAs are intronless transcripts, thereby precluding any possibility for distinction between the amplification products derived from cDNA and the genomic DNA ones. Currently, we are discussing alternative experimental approaches to circumvent the abovementioned technical difficulties.

5.3 SINE RNAs bind only a particular subset of dsRBPs

Previously, it has been observed that the transgenic *A. thaliana* lines that overexpress exogenous SINE RNAs, abundant highly structured retroelements consisting of three stem-loops with mismatches and bulges, exhibit phenotypic aberrations similar to *hyl1* and *drb4* mutants, deficient in proteins essential for RNAi in plants (Pouch-Pelissier et al., 2008). Thus, it is easy to imagine that SINE RNAs, bearing high structural similarity to microRNA intermediates, serve as modulators of RNA interference mechanism in plants. This is why we investigated *in vitro* interactions of SINE RNAs with different RNAi-related dsRBPs . We have tested if the affinity of dsRBD-containing proteins for this group of imperfect dsRNAs was compatible with their *in vivo* preferences and obtained the results presented below.

5.3.1 SINE RNA can bind HYL1 but not DRB4

In our *in vitro* experiments, HYL1 binds efficiently different imperfect double-stranded RNA molecules, including SB1 and SB2 SINE RNAs, while DRB4 only binds perfect RNA duplexes. These results are fully compatible with the known natural substrates of these two proteins. *In vivo*, HYL1 is known to interact with pri- and pre-miRNAs, which are organized as stem-loops containing mismatches and bulges. On the other hand, DRB4 binds perfect linear RNA duplexes formed by the action of RDR6 on a single stranded primary transcript (Hiraguri et al., 2005). The fact that HYL1 does not play a major role in double-stranded RNA-induced posttranscriptional gene silencing (PTGS) (Vazquez et al., 2004) further suggests that, *in vivo*, HYL1 preferentially interacts with imperfect double-stranded pri- and pre-miRNAs and not perfect double-stranded PTGS precursors. Based on our RNase V1 mapping results, the basis of this selectivity could be the capacity of HYL1 to interact with single-stranded RNA regions. Indeed, the binding specificity of other eukaryotic dsRBPs, such as ADARs and Staufen, was shown to depend on their ability to interact with single-stranded RNA loops (Ramos et al., 2000; Stefl et al., 2005). It is, therefore, easily concievable that RNA binding specificities of HYL1 are much different from DRB4, determining their different set of interaction partners *in vivo*.

5.3.2 Chaperoning-like activity of HYL1

We also observed that upon binding RNA HYL1 has the general capacity to force single-stranded regions to adopt a more organized, helix-like configuration. *In vivo*, HYL1 mainly is involved in promoting processing steps from pri-miRNA to pre-miRNA in association with DCL1, another dsRBP (Song et al., 2007; Wu et al., 2007). HYL1 also influences the cleavage positioning of DCL1 on the pre-miRNA to generate the mature miRNA (Kurihara et al., 2006). Consequently, in the *hyl1-2* null mutant, pri-miRNAs accumulate and misplaced cleavages of pre-miRNAs were observed in some cases. However, HYL1 is not fully necessary for plant miRNA processing by DCL1 because the *hyl1-2* mutant retains some ability to accumulate wild type miRNAs, although the accumulation level is reduced. Also, this reduction is variable depending on the different miRNAs (Han et al., 2004; Vazquez et al., 2004). HYL1 may therefore promote, to variable extents, the processing activity of DCL1. Based on our observations, one way HYL1 could do this is by inducing a conformational change in the RNA structure, forcing key single stranded regions to adopt organized, helix-like, configurations, including non Watson-Crick base pairing. This could in turn be important for promoting the cleavage activity of DCL1 on pri-miRNAs or for helping to precisely define the cleavage site on pre-miRNAs to generate mature products. It remains to be determined whether the chaperoning-like activity of HYL1 is important for miRNA production.

5.3.3 Possible evolutionary consequence of the SINE RNA/dsRBP interaction

Related structural domains can be found in many SINE RNAs from mammals, fishes and plants, suggesting common selective constraints imposed at the SINE RNA structural level (Sun et al., 2007). Using a double-stranded RNA binding domain from *Xenopus laevis* ADAR1, it is demonstrated in this work that the plant SB1 RNA is capable of binding strongly divergent dsRBPs. Therefore, it is plausible that the common trend of structural evolution observed for tRNA-related SINE RNAs could result in similar constraints imposed by a subset of dsRBPs across eukaryote domain of life. This rigid RNA-protein recognition mode would then enforce SINE RNAs to keep intact their capacity to interact with some dsRBPs. This would in turn forge the SINE RNA structure and impose, as observed, a common evolutionary history for

most eukaryote tRNA-related SINEs. The reason why SINE RNA/dsRBP interaction would be under positive selective pressure is unclear, but precise and punctual expression of SINEs during a key development step or in a stress situation, could induce genetic and/or epigenetic variations and increase diversity and/or adaptability. It is interesting to note that SINE-specific expression in their natural host is highly regulated at the transcriptional and post-transcriptional levels by complex genetic and epigenetic processes (Pelissier et al., 2004; Schmid, 1998). SINEs are non-autonomous in their mobility and need the activation of an autonomous LINE partner to retrotranspose. Therefore, based on our results, we suggest that the main purpose of limiting SINE-specific transcription is not to prevent its mobility (the control of LINEs is sufficient to achieve this) but to preserve cell homeostasis by preventing SINE RNA to interact with a large subset of dsRBPs.

6 REFERENCES

Agrawal, R. and Stormo, G.D. (2005) Editing efficiency of a Drosophila gene correlates with a distant splice site selection. *Rna*, **11**, 563-566.

Alfonzo, J.D., Thiemann, O. and Simpson, L. (1997) The mechanism of U insertion/deletion RNA editing in kinetoplastid mitochondria. *Nucleic Acids Res*, **25**, 3751-3759.

Allen, T.A., Von Kaenel, S., Goodrich, J.A. and Kugel, J.F. (2004) The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. *Nat Struct Mol Biol*, **11**, 816-821.

Arnaud, P., Yukawa, Y., Lavie, L., Pelissier, T., Sugiura, M. and Deragon, J.M. (2001) Analysis of the SINE S1 Pol III promoter from Brassica; impact of methylation and influence of external sequences. *Plant J*, **26**, 295-305.

Athanasiadis, A., Rich, A. and Maas, S. (2004) Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS Biol*, **2**, e391.

Bass, B.L. (1997) RNA editing and hypermutation by adenosine deamination. *Trends Biochem Sci*, **22**, 157-162.

Bass, B.L. (2002) RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem*, **71**, 817-846.

Bass, B.L. and Weintraub, H. (1987) A developmentally regulated activity that unwinds RNA duplexes. *Cell*, **48**, 607-613.

Bass, B.L. and Weintraub, H. (1988) An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell*, **55**, 1089-1098.

Benne, R., Van den Burg, J., Brakenhoff, J.P., Sloof, P., Van Boom, J.H. and Tromp, M.C. (1986) Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell*, **46**, 819-826.

Berglund, J.A., Chua, K., Abovich, N., Reed, R. and Rosbash, M. (1997) The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell*, **89**, 781-787.

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

Bjork, G.R., Ericson, J.U., Gustafsson, C.E., Hagervall, T.G., Jonsson, Y.H. and Wikstrom, P.M. (1987) Transfer RNA modification. *Annu Rev Biochem*, **56**, 263-287.

Blow, M.J., Grocock, R.J., van Dongen, S., Enright, A.J., Dicks, E., Futreal, P.A., Wooster, R. and Stratton, M.R. (2006) RNA editing of human microRNAs. *Genome Biol*, **7**, R27.

Bransteitter, R., Pham, P., Scharff, M.D. and Goodman, M.F. (2003) Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc Natl Acad Sci U S A*, **100**, 4102-4107.

Bratt, E. and Ohman, M. (2003) Coordination of editing and splicing of glutamate receptor pre-mRNA. *Rna*, **9**, 309-318.

Brown, B.A., 2nd, Lowenhaupt, K., Wilbert, C.M., Hanlon, E.B. and Rich, A. (2000) The zalpha domain of the editing enzyme dsRNA adenosine deaminase binds left-handed Z-RNA as well as Z-DNA. *Proc Natl Acad Sci U S A*, **97**, 13532-13536.

Brownawell, A.M. and Macara, I.G. (2002) Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins. *J Cell Biol*, **156**, 53-64.

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

Burns, C.M., Chu, H., Rueter, S.M., Hutchinson, L.K., Canton, H., Sanders-Bush, E. and Emeson, R.B. (1997) Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature*, **387**, 303-308.

Calado, A., Treichel, N., Muller, E.C., Otto, A. and Kutay, U. (2002) Exportin-5-mediated nuclear export of eukaryotic elongation factor 1A and tRNA. *Embo J*, **21**, 6216-6224.

Caputi, M. and Zahler, A.M. (2001) Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family. *J Biol Chem*, **276**, 43850-43859.

Cattaneo, R., Schmid, A., Eschle, D., Baczko, K., ter Meulen, V. and Billeter, M.A. (1988) Biased hypermutation and other genetic changes in defective measles viruses in human brain infections. *Cell*, **55**, 255-265.

Chan, L. (1992) Apolipoprotein B, the major protein component of triglyceride-rich and low density lipoproteins. *J Biol Chem*, **267**, 25621-25624.

Chen, C.D., Kobayashi, R. and Helfman, D.M. (1999) Binding of hnRNP H to an exonic splicing silencer is involved in the regulation of alternative splicing of the rat beta-tropomyosin gene. *Genes Dev*, **13**, 593-606.

Chen, C.X., Cho, D.S., Wang, Q., Lai, F., Carter, K.C. and Nishikura, K. (2000) A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *Rna*, **6**, 755-767.

Chen, S.H., Habib, G., Yang, C.Y., Gu, Z.W., Lee, B.R., Weng, S.A., Silberman, S.R., Cai, S.J., Deslypere, J.P., Rosseneu, M. and et al. (1987) Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in-frame stop codon. *Science*, **238**, 363-366.

Chen, T., Brownawell, A.M. and Macara, I.G. (2004) Nucleocytoplasmic shuttling of JAZ, a new cargo protein for exportin-5. *Mol Cell Biol*, **24**, 6608-6619.

Curtin, S.J., Watson, J.M., Smith, N.A., Eamens, A.L., Blanchard, C.L. and Waterhouse, P.M. (2008) The roles of plant dsRNA-binding proteins in RNAi-like pathways. *FEBS Lett*, **582**, 2753-2760.

David, C.J. and Manley, J.L. (2008) The search for alternative splicing regulators: new approaches offer a path to a splicing code. *Genes Dev*, **22**, 279-285.

Dawson, T.R., Sansam, C.L. and Emeson, R.B. (2004) Structure and sequence determinants required for the RNA editing of ADAR2 substrates. *J Biol Chem*, **279**, 4941-4951.

Decatur, W.A. and Fournier, M.J. (2003) RNA-guided nucleotide modification of ribosomal and other RNAs. *J Biol Chem*, **278**, 695-698.

Del Gatto-Konczak, F., Olive, M., Gesnel, M.C. and Breathnach, R. (1999) hnRNP A1 recruited to an exon in vivo can function as an exon splicing silencer. *Mol Cell Biol*, **19**, 251-260.

Dewannieux, M., Esnault, C. and Heidmann, T. (2003) LINE-mediated retrotransposition of marked Alu sequences. *Nat Genet*, **35**, 41-48.

Donahue, J.P., Vetter, M.L., Mukhtar, N.A. and D'Aquila, R.T. (2008) The HIV-1 Vif PPLP motif is necessary for human APOBEC3G binding and degradation. *Virology*, **377**, 49-53.

Doria, M., Neri, F., Gallo, A., Farace, M.G. and Michienzi, A. (2009) Editing of HIV-1 RNA by the double-stranded RNA deaminase ADAR1 stimulates viral infection. *Nucleic Acids Res*.

Dreyfuss, G., Adam, S.A. and Choi, Y.D. (1984) Physical change in cytoplasmic messenger ribonucleoproteins in cells treated with inhibitors of mRNA transcription. *Mol Cell Biol*, **4**, 415-423.

Eckmann, C.R. and Jantsch, M.F. (1999) The RNA-editing enzyme ADAR1 is localized to the nascent ribonucleoprotein matrix on Xenopus lampbrush chromosomes but specifically associates with an atypical loop. *J Cell Biol*, **144**, 603-615.

Espinoza, C.A., Allen, T.A., Hieb, A.R., Kugel, J.F. and Goodrich, J.A. (2004) B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. *Nat Struct Mol Biol*, **11**, 822-829.

Estevez, A.M. and Simpson, L. (1999) Uridine insertion/deletion RNA editing in trypanosome mitochondria--a review. *Gene*, **240**, 247-260.

Fairbrother, W.G. and Chasin, L.A. (2000) Human genomic sequences that inhibit splicing. *Mol Cell Biol*, **20**, 6816-6825.

Feagin, J.E., Abraham, J.M. and Stuart, K. (1988) Extensive editing of the cytochrome c oxidase III transcript in Trypanosoma brucei. *Cell*, **53**, 413-422.

Flomen, R., Knight, J., Sham, P., Kerwin, R. and Makoff, A. (2004) Evidence that RNA editing modulates splice site selection in the 5-HT2C receptor gene. *Nucleic Acids Res*, **32**, 2113-2122.

Fritz, J., Strehblow, A., Taschner, A., Schopoff, S., Pasierbek, P. and Jantsch, M.F. (2009) RNA-regulated interaction of transportin-1 and exportin-5 with the double-stranded RNA-binding domain regulates nucleocytoplasmic shuttling of ADAR1. *Mol Cell Biol*, **29**, 1487-1497.

Gallo, A., Keegan, L.P., Ring, G.M. and O'Connell, M.A. (2003) An ADAR that edits transcripts encoding ion channel subunits functions as a dimer. *Embo J*, **22**, 3421-3430.

Gallo, A., Thomson, E., Brindle, J., O'Connell, M.A. and Keegan, L.P. (2002) Micro-processing events in mRNAs identified by DHPLC analysis. *Nucleic Acids Res*, **30**, 3945-3953.

George, C.X. and Samuel, C.E. (1999) Human RNA-specific adenosine deaminase ADAR1 transcripts possess alternative exon 1 structures that initiate from different promoters, one constitutively active and the other interferon inducible. *Proc Natl Acad Sci U S A*, **96**, 4621-4626.

Gerber, A., Grosjean, H., Melcher, T. and Keller, W. (1998) Tad1p, a yeast tRNA-specific adenosine deaminase, is related to the mammalian pre-mRNA editing enzymes ADAR1 and ADAR2. *Embo J*, **17**, 4780-4789.

Gerber, A.P. and Keller, W. (1999) An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science*, **286**, 1146-1149.

Gerber, A.P. and Keller, W. (2001) RNA editing by base deamination: more enzymes, more targets, new mysteries. *Trends Biochem Sci*, **26**, 376-384.

Gott, J.M. and Emeson, R.B. (2000) Functions and mechanisms of RNA editing. *Annu Rev Genet*, **34**, 499-531.

Grabowski, P.J. (2004) A molecular code for splicing silencing: configurations of guanosine-rich motifs. *Biochem Soc Trans*, **32**, 924-927.

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

Greger, I.H. and Esteban, J.A. (2007) AMPA receptor biogenesis and trafficking. *Curr Opin Neurobiol*, **17**, 289-297.

Greger, I.H., Khatri, L., Kong, X. and Ziff, E.B. (2003) AMPA receptor tetramerization is mediated by Q/R editing. *Neuron*, **40**, 763-774.

Greger, I.H., Khatri, L. and Ziff, E.B. (2002) RNA editing at arg607 controls AMPA receptor exit from the endoplasmic reticulum. *Neuron*, **34**, 759-772.

Grosjean, H., Auxilien, S., Constantinesco, F., Simon, C., Corda, Y., Becker, H.F., Foiret, D., Morin, A., Jin, Y.X., Fournier, M. and Fourrey, J.L. (1996) Enzymatic conversion of adenosine to inosine and to N1-methylinosine in transfer RNAs: a review. *Biochimie*, **78**, 488-501.

Gurevich, I., Tamir, H., Arango, V., Dwork, A.J., Mann, J.J. and Schmauss, C. (2002) Altered editing of serotonin 2C receptor pre-mRNA in the prefrontal cortex of depressed suicide victims. *Neuron*, **34**, 349-356.

Guthrie, C. and Patterson, B. (1988) Spliceosomal snRNAs. *Annu Rev Genet*, **22**, 387-419.

Gwizdek, C., Ossareh-Nazari, B., Brownawell, A.M., Doglio, A., Bertrand, E., Macara, I.G. and Dargemont, C. (2003) Exportin-5 mediates nuclear export of minihelix-containing RNAs. *J Biol Chem*, **278**, 5505-5508.

Gwizdek, C., Ossareh-Nazari, B., Brownawell, A.M., Evers, S., Macara, I.G. and Dargemont, C. (2004) Minihelix-containing RNAs mediate exportin-5-dependent nuclear export of the double-stranded RNA-binding protein ILF3. *J Biol Chem*, **279**, 884-891.

Hallegger, M., Taschner, A. and Jantsch, M.F. (2006) RNA aptamers binding the double-stranded RNA-binding domain. *Rna*, **12**, 1993-2004.

Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T. and Kim, V.N. (2006) Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell*, **125**, 887-901.

Han, M.H., Goud, S., Song, L. and Fedoroff, N. (2004) The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proc Natl Acad Sci U S A*, **101**, 1093-1098.

Harris, R.S. and Liddament, M.T. (2004) Retroviral restriction by APOBEC proteins. *Nat Rev Immunol*, **4**, 868-877.

Hartner, J.C., Schmittwolf, C., Kispert, A., Muller, A.M., Higuchi, M. and Seeburg, P.H. (2004) Liver disintegration in the mouse embryo caused by deficiency in the RNA-editing enzyme ADAR1. *J Biol Chem*, **279**, 4894-4902.

Hartner, J.C., Walkley, C.R., Lu, J. and Orkin, S.H. (2009) ADAR1 is essential for the maintenance of hematopoiesis and suppression of interferon signaling. *Nat Immunol*, **10**, 109-115.

Hasler, J. and Strub, K. (2006) Alu elements as regulators of gene expression. *Nucleic Acids Res*, **34**, 5491-5497.

Heinrichs, V., Bach, M., Winkelmann, G. and Luhrmann, R. (1990) U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5' splice site. *Science*, **247**, 69-72.

Herbert, A., Alfken, J., Kim, Y.G., Mian, I.S., Nishikura, K. and Rich, A. (1997) A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. *Proc Natl Acad Sci U S A*, **94**, 8421-8426.

Herbert, A. and Rich, A. (2001) The role of binding domains for dsRNA and Z-DNA in the in vivo editing of minimal substrates by ADAR1. *Proc Natl Acad Sci U S A*, **98**, 12132-12137.

Herz, J. and Willnow, T.E. (1995) Lipoprotein and receptor interactions in vivo. *Curr Opin Lipidol*, **6**, 97-103.

Higuchi, M., Maas, S., Single, F.N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R. and Seeburg, P.H. (2000) Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature*, **406**, 78-81.

Higuchi, M., Single, F.N., Kohler, M., Sommer, B., Sprengel, R. and Seeburg, P.H. (1993) RNA editing of AMPA receptor subunit GluR-B: a base-paired intron-exon structure determines position and efficiency. *Cell*, **75**, 1361-1370.

Hiraguri, A., Itoh, R., Kondo, N., Nomura, Y., Aizawa, D., Murai, Y., Koiwa, H., Seki, M., Shinozaki, K. and Fukuhara, T. (2005) Specific interactions between Dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in Arabidopsis thaliana. *Plant Mol Biol*, **57**, 173-188.

Hoopengardner, B., Bhalla, T., Staber, C. and Reenan, R. (2003) Nervous system targets of RNA editing identified by comparative genomics. *Science*, **301**, 832-836.

Hough, R.F. and Bass, B.L. (1994) Purification of the Xenopus laevis double-stranded RNA adenosine deaminase. *J Biol Chem*, **269**, 9933-9939.

Hough, R.F. and Bass, B.L. (1997) Analysis of Xenopus dsRNA adenosine deaminase cDNAs reveals similarities to DNA methyltransferases. *Rna*, **3**, 356-370.

Hough, R.F., Lingam, A.T. and Bass, B.L. (1999) Caenorhabditis elegans mRNAs that encode a protein similar to ADARs derive from an operon containing six genes. *Nucleic Acids Res*, **27**, 3424-3432.

Hoyer, D., Clarke, D.E., Fozard, J.R., Hartig, P.R., Martin, G.R., Mylecharane, E.J., Saxena, P.R. and Humphrey, P.P. (1994) International Union of Pharmacology classification of receptors for 5-hydroxytryptamine (Serotonin). *Pharmacol Rev*, **46**, 157-203.

Kallman, A.M., Sahlin, M. and Ohman, M. (2003) ADAR2 A-->I editing: site selectivity and editing efficiency are separate events. *Nucleic Acids Res*, **31**, 4874-4881.

Kawahara, Y., Grimberg, A., Teegarden, S., Mombereau, C., Liu, S., Bale, T.L., Blendy, J.A. and Nishikura, K. (2008) Dysregulated editing of serotonin 2C receptor mRNAs results in energy dissipation and loss of fat mass. *J Neurosci*, **28**, 12834-12844.

Kawahara, Y., Ito, K., Sun, H., Aizawa, H., Kanazawa, I. and Kwak, S. (2004) Glutamate receptors: RNA editing and death of motor neurons. *Nature*, **427**, 801.

Kawahara, Y., Zinshteyn, B., Chendrimada, T.P., Shiekhattar, R. and Nishikura, K. (2007a) RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. *EMBO Rep*, **8**, 763-769.

Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A.G. and Nishikura, K. (2007b) Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science*, **315**, 1137-1140.

Keegan, L.P., Brindle, J., Gallo, A., Leroy, A., Reenan, R.A. and O'Connell, M.A. (2005) Tuning of RNA editing by ADAR is required in Drosophila. *Embo J*, **24**, 2183-2193.

Keegan, L.P., Gallo, A. and O'Connell, M.A. (2001) The many roles of an RNA editor. *Nat Rev Genet*, **2**, 869-878.

Keegan, L.P., Gerber, A.P., Brindle, J., Leemans, R., Gallo, A., Keller, W. and O'Connell, M.A. (2000) The properties of a tRNA-specific adenosine deaminase from Drosophila melanogaster support an evolutionary link between pre-mRNA editing and tRNA modification. *Mol Cell Biol*, **20**, 825-833.

Keegan, L.P., Leroy, A., Sproul, D. and O'Connell, M.A. (2004) Adenosine deaminases acting on RNA (ADARs): RNA-editing enzymes. *Genome Biol*, **5**, 209.

Kim, U., Garner, T.L., Sanford, T., Speicher, D., Murray, J.M. and Nishikura, K. (1994a) Purification and characterization of double-stranded RNA adenosine deaminase from bovine nuclear extracts. *J Biol Chem*, **269**, 13480-13489.

Kim, U., Wang, Y., Sanford, T., Zeng, Y. and Nishikura, K. (1994b) Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing. *Proc Natl Acad Sci U S A*, **91**, 11457-11461.

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

Konarska, M.M. and Sharp, P.A. (1987) Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell*, **49**, 763-774.

Krol, A., Westhof, E., Bach, M., Luhrmann, R., Ebel, J.P. and Carbon, P. (1990) Solution structure of human U1 snRNA. Derivation of a possible three-dimensional model. *Nucleic Acids Res*, **18**, 3803-3811.

Kurihara, Y., Takashi, Y. and Watanabe, Y. (2006) The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. *Rna*, **12**, 206-212.

Kwak, S. and Kawahara, Y. (2005) Deficient RNA editing of GluR2 and neuronal death in amyotropic lateral sclerosis. *J Mol Med*, **83**, 110-120.

Lai, F., Drakas, R. and Nishikura, K. (1995) Mutagenic analysis of double-stranded RNA adenosine deaminase, a candidate enzyme for RNA editing of glutamate-gated ion channel transcripts. *J Biol Chem*, **270**, 17098-17105.

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

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S. and Kim, V.N. (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature*, **425**, 415-419.

Lehmann, K.A. and Bass, B.L. (2000) Double-stranded RNA adenosine deaminases ADAR1 and ADAR2 have overlapping specificities. *Biochemistry*, **39**, 12875-12884.

Lenoir, A., Lavie, L., Prieto, J.L., Goubely, C., Cote, J.C., Pelissier, T. and Deragon, J.M. (2001) The evolutionary origin and genomic organization of SINEs in Arabidopsis thaliana. *Mol Biol Evol*, **18**, 2315-2322.

Lev-Maor, G., Sorek, R., Levanon, E.Y., Paz, N., Eisenberg, E. and Ast, G. (2007) RNA-editing-mediated exon evolution. *Genome Biol*, **8**, R29.

Levanon, E.Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., Shemesh, R., Fligelman, Z.Y., Shoshan, A., Pollock, S.R., Sztybel, D., Olshansky, M., Rechavi, G. and Jantsch, M.F. (2004) Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat Biotechnol*, **22**, 1001-1005.

Levanon, E.Y., Hallegger, M., Kinar, Y., Shemesh, R., Djinovic-Carugo, K., Rechavi, G., Jantsch, M.F. and Eisenberg, E. (2005) Evolutionarily conserved human targets of adenosine to inosine RNA editing. *Nucleic Acids Res*, **33**, 1162-1168.

Li, J.B., Levanon, E.Y., Yoon, J.K., Aach, J., Xie, B., Leproust, E., Zhang, K., Gao, Y. and Church, G.M. (2009) Genome-wide identification of human RNA editing sites by parallel DNA capturing and sequencing. *Science*, **324**, 1210-1213.

Liu, Y., George, C.X., Patterson, J.B. and Samuel, C.E. (1997) Functionally distinct double-stranded RNA-binding domains associated with alternative splice site variants of the interferon-inducible double-stranded RNA-specific adenosine deaminase. *J Biol Chem*, **272**, 4419-4428.

Liu, Y., Lei, M. and Samuel, C.E. (2000) Chimeric double-stranded RNA-specific adenosine deaminase ADAR1 proteins reveal functional selectivity of double-

stranded RNA-binding domains from ADAR1 and protein kinase PKR. *Proc Natl Acad Sci U S A*, **97**, 12541-12546.

Lomeli, H., Mosbacher, J., Melcher, T., Hoger, T., Geiger, J.R., Kuner, T., Monyer, H., Higuchi, M., Bach, A. and Seeburg, P.H. (1994) Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. *Science*, **266**, 1709-1713.

Lowman, H.B. and Draper, D.E. (1986) On the recognition of helical RNA by cobra venom V1 nuclease. *J Biol Chem*, **261**, 5396-5403.

Lu, C. and Fedoroff, N. (2000) A mutation in the Arabidopsis HYL1 gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin, and cytokinin. *Plant Cell*, **12**, 2351-2366.

Luciano, D.J., Mirsky, H., Vendetti, N.J. and Maas, S. (2004) RNA editing of a miRNA precursor. *Rna*, **10**, 1174-1177.

Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E. and Kutay, U. (2004) Nuclear export of microRNA precursors. *Science*, **303**, 95-98.

Ma, E., Gu, X.Q., Wu, X., Xu, T. and Haddad, G.G. (2001) Mutation in pre-mRNA adenosine deaminase markedly attenuates neuronal tolerance to O2 deprivation in Drosophila melanogaster. *J Clin Invest*, **107**, 685-693.

Maas, S., Gerber, A.P. and Rich, A. (1999) Identification and characterization of a human tRNA-specific adenosine deaminase related to the ADAR family of pre-mRNA editing enzymes. *Proc Natl Acad Sci U S A*, **96**, 8895-8900.

Maas, S., Kawahara, Y., Tamburro, K.M. and Nishikura, K. (2006) A-to-I RNA editing and human disease. *RNA Biol*, **3**, 1-9.

Maas, S., Kim, Y.G. and Rich, A. (2000) Sequence, genomic organization and functional expression of the murine tRNA-specific adenosine deaminase ADAT1. *Gene*, **243**, 59-66.

Maas, S., Patt, S., Schrey, M. and Rich, A. (2001) Underediting of glutamate receptor GluR-B mRNA in malignant gliomas. *Proc Natl Acad Sci U S A*, **98**, 14687-14692.

Maas, S. and Rich, A. (2000) Changing genetic information through RNA editing. *Bioessays*, **22**, 790-802.

Macbeth, M.R., Schubert, H.L., Vandemark, A.P., Lingam, A.T., Hill, C.P. and Bass, B.L. (2005) Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science*, **309**, 1534-1539.

Macchi, P., Brownawell, A.M., Grunewald, B., DesGroseillers, L., Macara, I.G. and Kiebler, M.A. (2004) The brain-specific double-stranded RNA-binding protein Staufen2: nucleolar accumulation and isoform-specific exportin-5-dependent export. *J Biol Chem*, **279**, 31440-31444.

Makarov, E.M., Makarova, O.V., Urlaub, H., Gentzel, M., Will, C.L., Wilm, M. and Luhrmann, R. (2002) Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science*, **298**, 2205-2208.

Mallory, A.C. and Vaucheret, H. (2006) Functions of microRNAs and related small RNAs in plants. *Nat Genet*, **38 Suppl**, S31-36.

Maniatis, T. and Reed, R. (1987) The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature*, **325**, 673-678.

Maslov, D.A. and Simpson, L. (1992) The polarity of editing within a multiple gRNA-mediated domain is due to formation of anchors for upstream gRNAs by downstream editing. *Cell*, **70**, 459-467.

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

Mehta, A., Kinter, M.T., Sherman, N.E. and Driscoll, D.M. (2000) Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA. *Mol Cell Biol*, **20**, 1846-1854.

Melcher, T., Maas, S., Herb, A., Sprengel, R., Seeburg, P.H. and Higuchi, M. (1996) A mammalian RNA editing enzyme. *Nature*, **379**, 460-464.

Mittaz, L., Antonarakis, S.E., Higuchi, M. and Scott, H.S. (1997) Localization of a novel human RNA-editing deaminase (hRED2 or ADARB2) to chromosome 10p15. *Hum Genet*, **100**, 398-400.

Miyamura, Y., Suzuki, T., Kono, M., Inagaki, K., Ito, S., Suzuki, N. and Tomita, Y. (2003) Mutations of the RNA-specific adenosine deaminase gene (DSRAD) are involved in dyschromatosis symmetrica hereditaria. *Am J Hum Genet*, **73**, 693-699.

Molineaux, S.M., Jessell, T.M., Axel, R. and Julius, D. (1989) 5-HT1c receptor is a prominent serotonin receptor subtype in the central nervous system. *Proc Natl Acad Sci U S A*, **86**, 6793-6797.

Morse, D.P., Aruscavage, P.J. and Bass, B.L. (2002) RNA hairpins in noncoding regions of human brain and Caenorhabditis elegans mRNA are edited by adenosine deaminases that act on RNA. *Proc Natl Acad Sci U S A*, **99**, 7906-7911.

Mount, S.M. (1982) A catalogue of splice junction sequences. *Nucleic Acids Res*, **10**, 459-472.

Murphy, J.T., Burgess, R.R., Dahlberg, J.E. and Lund, E. (1982) Transcription of a gene for human U1 small nuclear RNA. *Cell*, **29**, 265-274.

Nakazawa, Y., Hiraguri, A., Moriyama, H. and Fukuhara, T. (2007) The dsRNA-binding protein DRB4 interacts with the Dicer-like protein DCL4 in vivo and functions in the trans-acting siRNA pathway. *Plant Mol Biol*, **63**, 777-785.

Nelissen, R.L., Will, C.L., van Venrooij, W.J. and Luhrmann, R. (1994) The association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in part by common U snRNP proteins. *Embo J*, **13**, 4113-4125.

Nishikura, K. (2004) Editing the message from A to I. *Nat Biotechnol*, **22**, 962-963.

Nishikura, K. (2006) Editor meets silencer: crosstalk between RNA editing and RNA interference. *Nat Rev Mol Cell Biol*, **7**, 919-931.

Niswender, C.M., Sanders-Bush, E. and Emeson, R.B. (1998) Identification and characterization of RNA editing events within the 5-HT2C receptor. *Ann N Y Acad Sci*, **861**, 38-48.

O'Connell, M.A. and Keegan, L.P. (2006) Drosha versus ADAR: wrangling over pri-miRNA. *Nat Struct Mol Biol*, **13**, 3-4.

O'Connell, M.A. and Keller, W. (1994) Purification and properties of double-stranded RNA-specific adenosine deaminase from calf thymus. *Proc Natl Acad Sci U S A*, **91**, 10596-10600.

Ohlson, J., Pedersen, J.S., Haussler, D. and Ohman, M. (2007) Editing modifies the GABA(A) receptor subunit alpha3. *Rna*, **13**, 698-703.

Ohman, M. (2007) A-to-I editing challenger or ally to the microRNA process. *Biochimie*, **89**, 1171-1176.

Ohshima, K. and Okada, N. (2005) SINEs and LINEs: symbionts of eukaryotic genomes with a common tail. *Cytogenet Genome Res*, **110**, 475-490.

Oubridge, C., Ito, N., Evans, P.R., Teo, C.H. and Nagai, K. (1994) Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**, 432-438.

Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) Splicing of messenger RNA precursors. *Annu Rev Biochem*, **55**, 1119-1150.

Pagani, F. and Baralle, F.E. (2004) Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet*, **5**, 389-396.

Palladino, M.J., Keegan, L.P., O'Connell, M.A. and Reenan, R.A. (2000a) dADAR, a Drosophila double-stranded RNA-specific adenosine deaminase is highly developmentally regulated and is itself a target for RNA editing. *Rna*, **6**, 1004-1018.

Palladino, M.J., Keegan, L.P., O'Connell, M.A. and Reenan, R.A. (2000b) A-to-I pre-mRNA editing in Drosophila is primarily involved in adult nervous system function and integrity. *Cell*, **102**, 437-449.

Pan, Z.Q. and Prives, C. (1988) Assembly of functional U1 and U2 human-amphibian hybrid snRNPs in Xenopus laevis oocytes. *Science*, **241**, 1328-1331.

Parker, R., Siliciano, P.G. and Guthrie, C. (1987) Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell*, **49**, 229-239.

Patterson, J.B. and Samuel, C.E. (1995) Expression and regulation by interferon of a double-stranded-RNA-specific adenosine deaminase from human cells: evidence for two forms of the deaminase. *Mol Cell Biol*, **15**, 5376-5388.

Patton, D.E., Silva, T. and Bezanilla, F. (1997) RNA editing generates a diverse array of transcripts encoding squid Kv2 K+ channels with altered functional properties. *Neuron*, **19**, 711-722.

Patton, J.R. and Pederson, T. (1988) The Mr 70,000 protein of the U1 small nuclear ribonucleoprotein particle binds to the 5' stem-loop of U1 RNA and interacts with Sm domain proteins. *Proc Natl Acad Sci U S A*, **85**, 747-751.

Paul, M.S. and Bass, B.L. (1998) Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA. *Embo J*, **17**, 1120-1127.

Pelissier, T., Bousquet-Antonelli, C., Lavie, L. and Deragon, J.M. (2004) Synthesis and processing of tRNA-related SINE transcripts in Arabidopsis thaliana. *Nucleic Acids Res*, **32**, 3957-3966.

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

Polson, A.G., Bass, B.L. and Casey, J.L. (1996) RNA editing of hepatitis delta virus antigenome by dsRNA-adenosine deaminase. *Nature*, **380**, 454-456.

Polson, A.G., Crain, P.F., Pomerantz, S.C., McCloskey, J.A. and Bass, B.L. (1991) The mechanism of adenosine to inosine conversion by the double-stranded

RNA unwinding/modifying activity: a high-performance liquid chromatography-mass spectrometry analysis. *Biochemistry*, **30**, 11507-11514.

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

Pouch-Pelissier, M.N., Pelissier, T., Elmayan, T., Vaucheret, H., Boko, D., Jantsch, M.F. and Deragon, J.M. (2008) SINE RNA induces severe developmental defects in Arabidopsis thaliana and interacts with HYL1 (DRB1), a key member of the DCL1 complex. *PLoS Genet*, **4**, e1000096.

Poulsen, H., Nilsson, J., Damgaard, C.K., Egebjerg, J. and Kjems, J. (2001) CRM1 mediates the export of ADAR1 through a nuclear export signal within the Z-DNA binding domain. *Mol Cell Biol*, **21**, 7862-7871.

Powell, L.M., Wallis, S.C., Pease, R.J., Edwards, Y.H., Knott, T.J. and Scott, J. (1987) A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. *Cell*, **50**, 831-840.

Ramos, A., Grunert, S., Adams, J., Micklem, D.R., Proctor, M.R., Freund, S., Bycroft, M., St Johnston, D. and Varani, G. (2000) RNA recognition by a Staufen double-stranded RNA-binding domain. *Embo J*, **19**, 997-1009.

Rebagliati, M.R. and Melton, D.A. (1987) Antisense RNA injections in fertilized frog eggs reveal an RNA duplex unwinding activity. *Cell*, **48**, 599-605.

Riedmann, E.M., Schopoff, S., Hartner, J.C. and Jantsch, M.F. (2008) Specificity of ADAR-mediated RNA editing in newly identified targets. *Rna*, **14**, 1110-1118.

Rosbash, M. and Seraphin, B. (1991) Who's on first? The U1 snRNP-5' splice site interaction and splicing. *Trends Biochem Sci*, **16**, 187-190.

Rubin, C.M., Kimura, R.H. and Schmid, C.W. (2002) Selective stimulation of translational expression by Alu RNA. *Nucleic Acids Res*, **30**, 3253-3261.

Rueter, S.M., Dawson, T.R. and Emeson, R.B. (1999) Regulation of alternative splicing by RNA editing. *Nature*, **399**, 75-80.

Ryu, W.S., Bayer, M. and Taylor, J. (1992) Assembly of hepatitis delta virus particles. *J Virol*, **66**, 2310-2315.

Scadden, A.D. (2005) The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage. *Nat Struct Mol Biol*, **12**, 489-496.

Scadden, A.D. and Smith, C.W. (2001a) RNAi is antagonized by A-->I hyper-editing. *EMBO Rep*, **2**, 1107-1111.

Scadden, A.D. and Smith, C.W. (2001b) Specific cleavage of hyper-edited dsRNAs. *Embo J*, **20**, 4243-4252.

Schade, M., Turner, C.J., Kuhne, R., Schmieder, P., Lowenhaupt, K., Herbert, A., Rich, A. and Oschkinat, H. (1999) The solution structure of the Zalpha domain of the human RNA editing enzyme ADAR1 reveals a prepositioned binding surface for Z-DNA. *Proc Natl Acad Sci U S A*, **96**, 12465-12470.

Schmid, C.W. (1998) Does SINE evolution preclude Alu function? *Nucleic Acids Res*, **26**, 4541-4550.

Schoft, V.K., Schopoff, S. and Jantsch, M.F. (2007) Regulation of glutamate receptor B pre-mRNA splicing by RNA editing. *Nucleic Acids Res*, **35**, 3723-3732.

Schwartz, T., Rould, M.A., Lowenhaupt, K., Herbert, A. and Rich, A. (1999) Crystal structure of the Zalpha domain of the human editing enzyme ADAR1 bound to left-handed Z-DNA. *Science*, **284**, 1841-1845.

Seeburg, P.H., Higuchi, M. and Sprengel, R. (1998) RNA editing of brain glutamate receptor channels: mechanism and physiology. *Brain Res Brain Res Rev*, **26**, 217-229.

Sen, G.C. (2001) Viruses and interferons. *Annu Rev Microbiol*, **55**, 255-281.

Sharp, P.A. (1987) Splicing of messenger RNA precursors. *Science*, **235**, 766-771.

Sheehy, A.M., Gaddis, N.C., Choi, J.D. and Malim, M.H. (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature*, **418**, 646-650.

Sheppard, R.D., Raine, C.S., Bornstein, M.B. and Udem, S.A. (1985) Measles virus matrix protein synthesized in a subacute sclerosing panencephalitis cell line. *Science*, **228**, 1219-1221.

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

Sommer, B., Kohler, M., Sprengel, R. and Seeburg, P.H. (1991) RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell*, **67**, 11-19.

Song, L., Han, M.H., Lesicka, J. and Fedoroff, N. (2007) Arabidopsis primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. *Proc Natl Acad Sci U S A*, **104**, 5437-5442.

Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A. and Steinberg, S. (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res*, **26**, 148-153.

Stefl, R., Skrisovska, L. and Allain, F.H. (2005) RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Rep*, **6**, 33-38.

Stefl, R., Xu, M., Skrisovska, L., Emeson, R.B. and Allain, F.H. (2006) Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs. *Structure*, **14**, 345-355.

Stephens, O.M., Yi-Brunozzi, H.Y. and Beal, P.A. (2000) Analysis of the RNA-editing reaction of ADAR2 with structural and fluorescent analogues of the GluR-B R/G editing site. *Biochemistry*, **39**, 12243-12251.

Strehblow, A., Hallegger, M. and Jantsch, M.F. (2002) Nucleocytoplasmic distribution of human RNA-editing enzyme ADAR1 is modulated by double-stranded RNA-binding domains, a leucine-rich export signal, and a putative dimerization domain. *Mol Biol Cell*, **13**, 3822-3835.

Sun, F.J., Fleurdepine, S., Bousquet-Antonelli, C., Caetano-Anolles, G. and Deragon, J.M. (2007) Common evolutionary trends for SINE RNA structures. *Trends Genet*, **23**, 26-33.

Tonkin, L.A. and Bass, B.L. (2003) Mutations in RNAi rescue aberrant chemotaxis of ADAR mutants. *Science*, **302**, 1725.

Tonkin, L.A., Saccomanno, L., Morse, D.P., Brodigan, T., Krause, M. and Bass, B.L. (2002) RNA editing by ADARs is important for normal behavior in Caenorhabditis elegans. *Embo J*, **21**, 6025-6035.

Urlaub, H., Raker, V.A., Kostka, S. and Luhrmann, R. (2001) Sm protein-Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure. *Embo J*, **20**, 187-196.

Vaucheret, H. (2005) MicroRNA-dependent trans-acting siRNA production. *Sci STKE*, **2005**, pe43.

Vazquez, F., Gasciolli, V., Crete, P. and Vaucheret, H. (2004) The nuclear dsRNA binding protein HYL1 is required for microRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Curr Biol*, **14**, 346-351.

Wagner, R.W., Smith, J.E., Cooperman, B.S. and Nishikura, K. (1989) A double-stranded RNA unwinding activity introduces structural alterations by means of adenosine to inosine conversions in mammalian cells and Xenopus eggs. *Proc Natl Acad Sci U S A*, **86**, 2647-2651.

Wang, Q., Miyakoda, M., Yang, W., Khillan, J., Stachura, D.L., Weiss, M.J. and Nishikura, K. (2004a) Stress-induced apoptosis associated with null mutation of ADAR1 RNA editing deaminase gene. *J Biol Chem*, **279**, 4952-4961.

Wang, Q., O'Brien, P.J., Chen, C.X., Cho, D.S., Murray, J.M. and Nishikura, K. (2000) Altered G protein-coupling functions of RNA editing isoform and splicing variant serotonin2C receptors. *J Neurochem*, **74**, 1290-1300.

Wang, Z. and Burge, C.B. (2008) Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *Rna*, **14**, 802-813.

Wang, Z., Rolish, M.E., Yeo, G., Tung, V., Mawson, M. and Burge, C.B. (2004b) Systematic identification and analysis of exonic splicing silencers. *Cell*, **119**, 831-845.

Wang, Z., Xiao, X., Van Nostrand, E. and Burge, C.B. (2006) General and specific functions of exonic splicing silencers in splicing control. *Mol Cell*, **23**, 61-70.

Will, C.L., Rumpler, S., Klein Gunnewiek, J., van Venrooij, W.J. and Luhrmann, R. (1996) In vitro reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E) spliceosomal complexes. *Nucleic Acids Res*, **24**, 4614-4623.

Wong, S.K., Sato, S. and Lazinski, D.W. (2001) Substrate recognition by ADAR1 and ADAR2. *Rna*, **7**, 846-858.

Wong, S.K., Sato, S. and Lazinski, D.W. (2003) Elevated activity of the large form of ADAR1 in vivo: Very efficient RNA editing occurs in the cytoplasm. *Rna*, **9**, 586-598.

Wu, F., Yu, L., Cao, W., Mao, Y., Liu, Z. and He, Y. (2007) The N-terminal double-stranded RNA binding domains of Arabidopsis HYPONASTIC LEAVES1 are sufficient for pre-microRNA processing. *Plant Cell*, **19**, 914-925.

Wu, S., Romfo, C.M., Nilsen, T.W. and Green, M.R. (1999) Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature*, **402**, 832-835.

Yang, W., Chendrimada, T.P., Wang, Q., Higuchi, M., Seeburg, P.H., Shiekhattar, R. and Nishikura, K. (2006) Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat Struct Mol Biol*, **13**, 13-21.

Yang, W., Wang, Q., Howell, K.L., Lee, J.T., Cho, D.S., Murray, J.M. and Nishikura, K. (2005) ADAR1 RNA deaminase limits short interfering RNA efficacy in mammalian cells. *J Biol Chem*, **280**, 3946-3953.

Yi-Brunozzi, H.Y., Stephens, O.M. and Beal, P.A. (2001) Conformational changes that occur during an RNA-editing adenosine deamination reaction. *J Biol Chem*, **276**, 37827-37833.

Yi, R., Qin, Y., Macara, I.G. and Cullen, B.R. (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev*, **17**, 3011-3016.

Zalfa, F., Giorgi, M., Primerano, B., Moro, A., Di Penta, A., Reis, S., Oostra, B. and Bagni, C. (2003) The fragile X syndrome protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses. *Cell*, **112**, 317-327.

Zamore, P.D., Patton, J.G. and Green, M.R. (1992) Cloning and domain structure of the mammalian splicing factor U2AF. *Nature*, **355**, 609-614.

Zhang, Z. and Carmichael, G.G. (2001) The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell*, **106**, 465-475.

Zhuang, Y. and Weiner, A.M. (1986) A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell*, **46**, 827-835.

Zhuang, Y. and Weiner, A.M. (1989) A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes Dev*, **3**, 1545-1552.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*, **31**, 3406-3415.

ACKNOWLEDGEMENTS

This is to express my deepest gratitude to all people who contributed to this work in one way or another...

My supervisor Michael Jantsch is thanked for offering me the opportunity to explore many different aspects of RNA biology, guiding me on the way and paying all the bills...

All the past and present colleagues and friends from MJ lab are forever remembered for making my stay there as pleasant as it was...

Reviewers of this thesis are acknowledged for taking the trouble of reading through this versatile work and evaluating it...

The members of the Department of Chromosome Biology have always been friendly, creating a nice working atmosphere...

Friends in Sinj, Zagreb and Vienna are thanked for all the extracurricular activities and an occasional shoulder to cry on...

Last but not least, I want to say a big thank you to my family for being there ever since the 23rd of July, 1980 (and 9 months before, obviously)...

HVALA-THANK YOU-DANKESCHÖN

CURRICULUM VITAE

PERSONAL DATA

Full name: Draško Boko
Date of birth: 23.07.1980.
Place of birth: Sinj, Croatia
Citizenship: Croatian
Present address: Gablenzgasse 33/27, A-1150, Vienna, Austria
E-mail: drasko.boko@univie.ac.at

EDUCATION

School education:

09/1986 – 06/1994: Primary school in Sinj, Croatia

09/1994 – 06/1998: Secondary school Gimnazija Dinka Šimunovića Sinj, Croatia

Higher education:

10/1998 – 07/2004: Undergraduate studies of molecular biology, Department of Biology, University of Zagreb, Croatia

11/2003 – 07/2004: Diploma thesis research, Department of Biochemistry, University of Zagreb, Croatia
Title of thesis: Amber suppression in bacterium Escherichia coli using archaeal orthologous seryl-tRNA synthetase/suppresor tRNA pair
Supervisor: Prof. Dr. Ivana Weygand-Đurašević

02/2005 – ongoing: Doctoral studies of molecular biology, University of Vienna, Austria

Institution: Department of Chromosome Biology, Max F. Perutz Laboratories, University of Vienna

Title: Role of RNA editing in RNA splicing and in nuclear export of microRNA precursors

Supervisor: ao Univ-Prof. Dr. Michael Jantsch

POSTERS/PRESENTATIONS:

09/2004 – 10/2004: Conference of the Croatian Society of Molecular Biology and Biochemistry (HDBMB), Bjelolasica, Croatia

Title of poster: Lesjak S, Boko D, Rokov-Plavec J, Weygand-Đurašević I: Amber suppression in bacterium *Escherichia coli* using archaeal orthologous seryl-tRNA synthetase/suppressor tRNA pair.

07/2008 – 08/2008: 13th Annual Meeting of the RNA Society, Berlin, Germany

Title of poster: Boko D and Jantsch M: Regulation of splicing by RNA editing in protein-coding targets.

PUBLICATIONS:

Pouch-Pelissier, M.N., Pelissier, T., Elmayan, T., Vaucheret, H., Boko, D., Jantsch, M.F. and Deragon, J.M. (2008) SINE RNA induces severe developmental defects in Arabidopsis thaliana and interacts with HYL1 (DRB1), a key member of the DCL1 complex. *PLoS Genet*, **4**, e1000096.

Lesjak, S., Boko D. and Weygand-Đurašević I. (2009) Seryl-tRNA synthetases from methanogenic archaea: suppression of bacterial amber mutation and heterologous toxicity. (submitted).