The Biochemical Analysis of the *Drosophila* RNA Editing Enzyme ADAR and its functional domains.

Gillian M. Ring

PhD Thesis

University of Edinburgh

April 2006

For Gran and Granda Campbell

"The great tragedy of Science – the slaying of a beautiful hypothesis by an ugly fact"

Thomas H. Huxley

English Biologist (1825 – 1895)

ABSTRACT

RNA editing is an mRNA processing event that involves the insertion, deletion or substitution of bases in the pre-mRNA molecule and has the potential to alter the genetic code resulting in a range of proteins with a different primary sequence than that suggested by the DNA. Its importance has only been fully realised now that the sequence of the human genome is being finalised and that the number of genes it contains are much lower than expected.

The adenosine deaminases acting on RNA or ADAR enzymes convert adenosine to inosine in dsRNA. They are members of the larger cytidine deaminase family and consist of a C-terminal catalytic domain and two or more double-stranded RNA binding motifs (dsRBMs). In *Drosophila melanogaster* there is only one *adar* gene. The *Drosophila* ADAR (dADAR) enzyme consists of two dsRBMs and a deaminase domain and is known to edit a number of transcripts encoding ion channels in the central nervous system. The generation of 3 different dADAR isoforms at varying stages of fly development enables enzyme activity to be regulated, with the adult form being the most active.

The ADAR family of enzymes is thought to have evolved from the ADAT (adenosine deaminase acting on tRNA) family that convert adenosine to inosine in tRNA, by adding dsRNA binding motifs to a deaminase domain. *Drosophila* ADAT1 is very homologous to the deaminase domain of the dADAR enzyme, both having 3 distinctive motifs containing residues thought to chelate zinc. Studies have shown that the deaminase domain of the ADAR enzymes is responsible for substrate specificity. I have cloned and purified the *Drosophila* deaminase domain and assayed its editing activity *in vivo* and *in vitro* on the dADAR self edit site, and on a dsRNA. The results suggest that the dADAR deaminase domain can edit dsRNA by itself. The

editing pattern on the dsRNA mirrors that seen with the full-length wild-type ADAR enzyme suggesting that site specificity is indeed conferred by the deaminase domain.

The absence of an alternatively spliced exon between the two dsRBMs of dADAR in adult flies results in higher editing activity when compared to the isoform containing the exon seen at earlier stages of development. A number of constructs containing the dsRBMs were successfully cloned and purified. NMR data of RBM1 confirms that the dsRBMs of dADAR are members of the well characterised family of dsRBMs with a typical α helix, 3 β sheets, α helix structure and each dsRBM has a distinct affinity for dsRNA substrate indicating possible different roles. Point mutations within dsRBM1 result in a loss of binding activity.

The self-editing site of dADAR is found within exon 7 of the pre-mRNA, 3′ to the second catalytic motif of the deaminase domain. Editing at this site results in a serine to glycine change in the translated protein and has been shown to reduce editing activity in all known substrates. This region of the protein is highly conserved throughout the ADARs and sequence comparisons with human ADAR2 at the DNA level show it has the potential to be edited at the same site as dADAR. When the same serine is mutated to glycine in human ADAR2 a decrease in activity is again seen suggesting that this serine residue plays a role in the active site of the enzyme.

My results confirm the hypothesis that the catalytic deaminase domain acquired dsRBMs, increasing the efficiency of the editing reaction on dsRNA. The dsRBMs bind non-specifically to dsRNA altering and stabilising the structure, allowing the catalytic domain easier access to the adenosine nucleotide to be edited. However, because the deaminase domain has activity on its own it too must play a role in binding to dsRNA substrates. The characteristics of a dsRNA substrate do not involve an easily delineated nucleotide sequence but may instead involve a more complex structural

motif. This is not recognized by the dsRBMs but by the deaminase domain itself, probably in a way similar to how the ADATs recognize the evolutionary conserved structure of tRNAs.

ACKNOWLEDGEMENTS

Well, the fact that I've actually managed to get to this point is in itself, quite miraculous! But here I am and I have many, many people to thank for it.

First of all, the lab; thanks to Mary for all her help, especially over the last few months; to Jim for the laughs, bramble wine and for always having a secret stash of essential reagents; to Angela for her help and tips in the lab, her advice and guidance and generally explosive personality; to Anne for the many papers you've ordered, cakes you've made and beds you've provided over the last while; and finally to Liam, for forever reminding me of thermal expansion and the gas law!!

To the rest of Chromosome Biology; thanks to Howard, Wendy and Mary T for advice and guidance; to all in the RNA office, Barbara, Gavin, Jeremy, Dasa and Demian for the giggles and gossip; and a special thanks to Niki for housing me! Also a special thanks to Sandy for contributing an artistic element to my figures.

And to the many friends I made in Edinburgh; Shelagh, Nic, Dexy, Heidi, George, Sue, Steph, Sev, NtB, Jeff, Ian, Jane, Sapna and Dunc; thanks for the cakes, castles, cups of tea, shoulders to cry on, and beds to sleep in.

And last but most certainly not least; a big thanks to my parents, Margaret and Kenny, and my brother, Christopher for always being there for me and for still being here.

DECLARATION

I declare that this thesis has been composed by me, and all of the work is my own unless otherwise stated.

Gillian Margaret Ring

April 2006

TABLE OF CONTENTS

TABLE OF FIGURES

ABBREVIATIONS

A	adenosine (purine) or alanine
aa	amino acid
ACF	APOBEC-1 complementation factor
ADAR	adenosine deaminase acting on dsRNA
ADAT	adenosine deaminase acting on tRNA
AID	activation-induced deaminase
Ala	alanine
ALS	Amyotrophic Lateral Sclerosis
apoB	Apolipoprotein B
APOBEC-1	apoB mRNA editing catalytic component 1
AML	acute myeloid leukaemia
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionate
ATP	adenosine triphosphate
AU	adenosine/uridine
BMGY	Buffered glycerol-complex medium
BMMY	Buffered methanol-complex medium
BSA	bovine serum albumin
C	cytosine (pyrimidine) or cytidine
$CaCl_2$	calcium chloride
cADAR1	constitutively expressed adenosine deaminase acting on dsRNA 1
CDA	cytidine deaminase
cDNA	complementary DNA
COSY	homonuclear correlation spectroscopy
C-terminal	carboxy terminal
CSR	class switch recombination
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
ddTTP	dideoxythymine triphosphate
dGTP	deoxyguanine triphosphate
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid

DNase	deoxyribonuclease
dNTP	deoxynucleoside triphosphate
Drosophila	*Drosophila melanogaster*
dsRBD	dsRNA binding domain
dsRBM	dsRNA binding motif
dsRNA	double strand ribonucleic acid
DTT	dithiothreitol
dTTP	deoxythymine triphosphate
dUTP	deoxyuridine triphosphate
E	glutamic acid
ECS	editing site complementary sequence
EDTA	ethylenediamine-N,N,N•,N'-tetra-acetic acid
EIAV	equine infectious anaemia virus
EST	expressed sequence tag
ETB	human endothelin B receptor gene
G	guanine (purine) or glycine
g	gravities
Glu-R	glutamate-gated receptor
H	histidine
HCl	hydrochloric acid
HDAg	hepatitis delta antigen
HDAg-L	hepatitis delta antigen, long form
HDAg-S	hepatitis delta antigen, short form
HDV	hepatitis delta virus
HIGM2	hyper-IgM syndrome (autosomal recessive form)
HIS	histidine
HSCR	hirschsprung disease
HSQC	heteronuclear single quantum correlation
I	inosine
IAA	iodoacetic acid
iADAR1	interferon – inducible adenosine deaminase acting on dsRNA 1
Ig	immunoglobulin
IGC	immunoglobulin gene conversion
IPTG	Isopropyl β-D-1-thiogalactopyranoside

K	lysine
kb	kilobase pairs of DNA
KCl	potassium chloride
kDa	kilo Daltons (molecular weight/10^3)
KH_2PO_4	potassium dihydrogen phosphate
LDL	low density lipoproteins
lnRNP	Large nuclear ribonucleoprotein
M	molar
MEF	mouse embryonic fibroblast
$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulphate
MIBE	measles inclusion body encephalitis
miRNA	micro RNA
MLV	murine leukaemia virus
mM	millimolar
mRNA	messenger RNA
MV	measles virus
NaCl	sodium chloride
Na_3PO_4	sodium phosphate
NAT1	novel APOBEC-1 target 1
NES	nuclear export signal
Ni^{2+}-NTA	nickel - nitrilotriacetic acid
NLS	nuclear localisation signal
nM	nanomolar
NMR	nuclear magnetic resonance
NOESY	nuclear overhauser effect spectroscopy
N-terminal	amino terminal
NTP	nucleoside triphosphate
OD	optical density
p	chromosome short arm
pH	
P	proline
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pgk-neo	phosphoglycerine kinase promoter driving the *neo* gene
PKR	double strand RNA activated protein

	kinase
PMSF	phenylmethanesulfonyl fluoride
q	chromosome long arm
Q	glutamine
R	arginine
RAG	rearrangement activating genes
RISC	RNAi induced silencing complex
RNA	ribonucleic acid
RNase	ribonuclease
RNAi	RNA interference
RRM	RNA recognition motif
rRNA	ribosomal RNA
S	serine
SAM	S-adenosylmethionine
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SH	sulfhydryl
SHM	somatic hypermutation
SLE	systemic lupus erythematosus
snoRNA	Small nucleolar RNA
SNP	small nucleotide polymorphism
snRNA	Small nuclear RNA
ssDNA	single strand deoxyribonucleic acid
SSPE	subacute sclerosing panencephalitis
ssRNA	single strand ribonucleic acid
T	thymidine (pyrimidine)
TBE	Tris-borate EDTA (90 mM Tris-borate, 2 mM EDTA)
TE	Tris, EDTA buffer (10 mM Tris, 1 mM EDTA, pH 8.0)
TOCSY	total correlation spectroscopy
TRBP	TAR RNA binding protein
tRNA	transfer ribonucleic acid
U	uridine
UDG	uracil-DNA glycosylase
UTR	untranslated region
UV	ultraviolet

V	valine
VIF	virion infectivity factor
VLDL	very low density lipoproteins
W	tryptophan
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
XlrbpA	*Xenopus laevis* RNA binding protein A
YNB	Yeast nitrogen base
YPD	Yeast extract peptone dextrose medium

1

INTRODUCTION

1.1 RNA EDITING

RNA editing is a term used to describe the structural alteration, insertion or deletion of nucleotides in RNA (Grosjean and Benne, 1998). If modification occurs in the messenger RNA it can result in the translation of a protein with a different primary sequence than that suggested by the DNA sequence of the gene. This process plays an important role in creating functional diversity in the protein products of gene expression.

RNA editing was first discovered in the mitochondria of primitive eukaryotes where the deletion or insertion of bases created new open reading frames (Benne et al., 1986). Mammalian editing is in the form of base alteration where a hydrolytic deamination of a cytidine or adenosine base results in uridine and inosine, respectively.

1.2 C-TO-U EDITING

C-to-U editing of the apolipoprotein B (apoB) transcript was the first example of RNA editing identified in mammals (Chen et al., 1987; Powell et al., 1987). Since this discovery, A-to-I RNA editing as well as C-to-U editing of DNA have been extensively studied.

1.2.1 C–to–U editing in RNA

1.2.1.1 The apoB editing machinery

Studies on cDNA clones showed a single C-to-U base change at C^{6666} of the apoB-100 mRNA resulted in an in-frame UAA stop codon replacing glutamate 2153 and resulting in the apoB-48 intestinal transcript (Chen et al., 1987; Powell et al., 1987) (Figure 1.1). This C-to-U change was later identified as an RNA editing event catalysed by the APOBEC-1 (apoB mRNA editing catalytic component 1) enzyme (Driscoll and Zhang, 1994; Lau et al., 1994; Teng et al., 1993). Further studies indicated that the human *APOBEC-1* gene is located on chromosome 12p13.1 – p13.2 and cDNA analysis predicts a 236 amino acid protein of approximately 27 kDa (Lau et al., 1994). The transcript is only found in the small intestine with the transcribed protein forming homo dimers but unable to edit the apoB transcript on its own (Lau et al., 1994; Navaratnam et al., 1993; Teng et al., 1993). Instead, an editing complex composed of an APOBEC-1 dimer and a 63.4 kDa auxiliary factor known as APOBEC-1 complementation factor (ACF) is the minimum requirement for apoB editing (Lellek et al., 2000; Mehta et al., 2000).

The APOBEC-1 protein is similar to the cytidine deaminase family especially in the region of its active site. Two conserved motifs found in this region, C/HAE and PCXXC, are involved in zinc chelation and proton shuttling during deamination and site-directed mutagenesis at these sites abolishes catalytic activity (Driscoll and Zhang, 1994; MacGinnitie et al., 1995). The RNA-binding activity of APOBEC-1 is relatively weak when compared to other cytidine deaminases and as yet, a distinct RNA binding domain has not been identified (Anant et al., 1995; Navaratnam et al., 1995).

A 65 kDa apoB editing auxiliary factor was purified and its structural characteristics analysed. Named APOBEC-1 complementation factor or ACF,

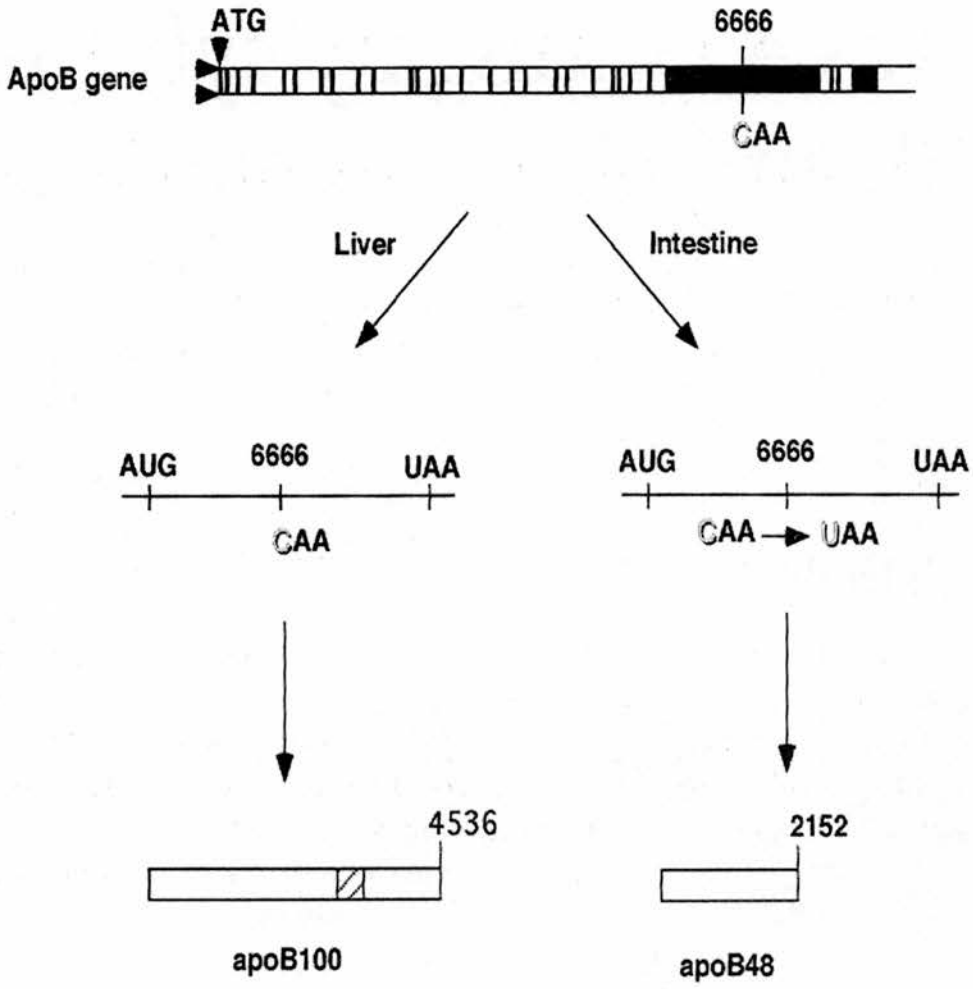


Figure 1.1 Cytidine deamination of apolipoptotein B.

A schematic of the human apoB gene. Concurrent with splicing and polyadenylation, C-6666 in the intestinal apoB mRNA undergoes site-specific deamination resulting in a U at this position. The proteins translated from the unedited and edited mRNAs are represented below.

(Taken from RNA Editing, 2000 Ed. B.L. Bass)

the protein contains three non-identical RNA recognition motifs (RRM) and is widely expressed in human tissues (Lellek et al., 2000; Mehta et al., 2000). Along with APOBEC-1, ACF comprises the minimal protein requirements for apoB editing and is thought to act as the RNA binding subunit of the editosome (Lellek et al., 2000; Mehta et al., 2000). More recently, ACF has been shown to be responsible for bringing APOBEC-1 from the cytoplasm into the nucleus where apoB RNA editing occurs (Blanc et al., 2003).

Other auxiliary factors have been identified and are believed to play a role in the editosome complex though as demonstrated only ACF and APOBEC-1 appear to be essential (Lellek et al., 2000; Schock et al., 1996).

1.2.1.2 Editing of the apoB transcript

The *apoB* gene is 43 kb in length and encodes 29 exons. Two forms of this gene exist, apoB100 (512 kDa) and apoB48 (241 kDa). ApoB100 is synthesised in the liver and is involved in the transport of cholesterol and triglycerides on very low density lipoproteins (VLDL). ApoB48 is found in the small intestine and is involved in the synthesis, assembly and secretion of chylomicrons. This shorter isoform is generated by the C-to-U editing event at cytidine 6666 (C^{6666}) which converts a glutamine (CAA) codon to a stop codon (UAA) in exon 26 (Innerarity et al., 1996).

The apoB mRNA sequences around C^{6666} are highly conserved across many species. Downstream of the editing site is an 11-nucleotide sequence called the "mooring" sequence (Shah et al., 1991). Point mutations in this sequence abolish editing (Driscoll et al., 1993; Shah et al., 1991) and the binding of the RRMs of ACF to the apoB mRNA are dependent on this sequence being intact (Mehta et al., 2000). The region between the edited C^{6666} and the mooring sequence is also important and is know as the "spacer" element. The length of this can range between 2 and 8 nucleotides, with the optimal length being 5 nucleotides (Driscoll et al., 1993).

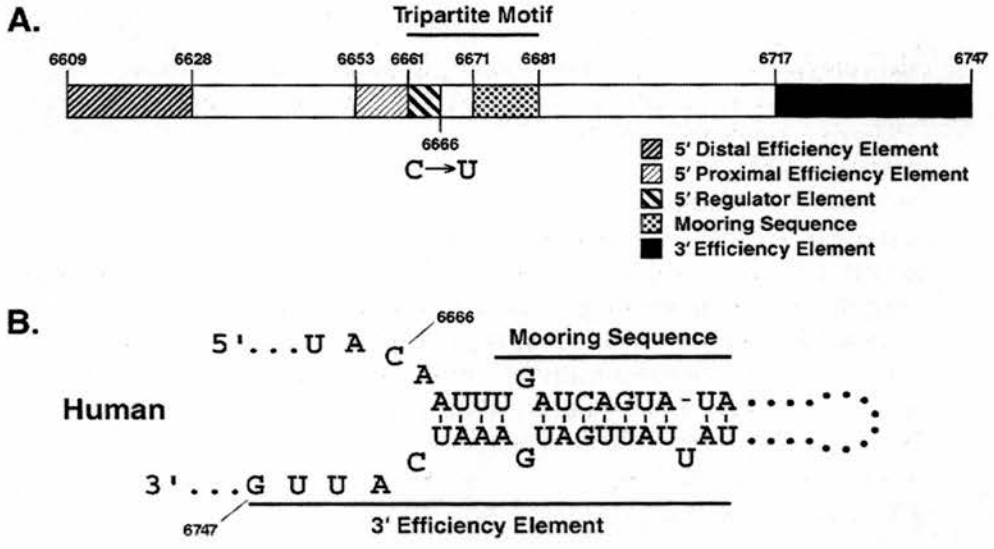


Figure 1.2 The apolipoprotein B substrate

A. Model of apoB mRNA in the region of the editing site showing different structural elements important for apoB mRNA editing.

B. Proposed secondary structure of human apoB mRNA.

(Taken from RNA Editing, 2000 Ed. B.L. Bass)

A 5 nucleotide region immediately 5′ of the editing site deemed the 5′ regulator element has been shown to play a significant role in editing efficiency although unlike the mooring sequence it is not essential (Backus and Smith, 1992). A number of other efficiency elements have been identified both 5′ and 3′ to the editing site and mutations in these regions reduce editing activity *in vitro* (Backus and Smith, 1991; Driscoll et al., 1993) (Figure 1.2).

1.2.1.3 Overexpression and deletions of APOBEC

Studies in low density lipoprotein (LDL) receptor deficient mice and rabbits demonstrated that an increase in apoB editing in the liver lowered LDL and VLDL levels in plasma (Teng et al., 1997). This suggested that APOBEC-1 may be a good candidate in gene therapy treatments of elevated LDL and VLDL levels which are often associated with coronary heart disease. However, overexpression of APOBEC-1 in the livers of mice and rabbits resulted in hepatocellular carcinomas. This was probably due to the indiscriminate editing of other transcripts involved in cell growth and differentiation (Yamanaka et al., 1995). Yamanaka *et al.* later identified a novel APOBEC-1 target, NAT1, which is highly conserved among species. This protein has homology to the carboxy-terminal portion of the eukaryotic translation initiation factor, eIF4G and is thought to be a translational repressor. When APOBEC-1 is overexpressed the NAT1 transcript is edited at several sites. This changes the primary sequence and generates stop codons resulting in considerably lower levels of NAT1 protein in the livers of mice overexpressing APOBEC-1 (Yamanaka et al., 1997). This promiscuous editing of NAT1 is thought to play a role in the oncogenic phenotype of these transgenic mice.

Hirano *et al.* and Morrison *et al.* generated *apobec-1* $^{-}/^{-}$ mice by targeted disruption of the gene (Hirano et al., 1996; Morrison et al., 1996). Homozygous mice were viable and fertile and displayed no obvious

phenotype other than increased levels of low density lipoprotein and reduced levels of high density lipoprotein. No apoB48 was detected indicating that the major role of APOBEC-1 is apo B RNA editing and the modulation of lipid transport (Hirano et al., 1996; Morrison et al., 1996; Nakamuta et al., 1996).

1.2.2 C–to–U editing in DNA

1.2.2.1 Antibody diversification

Antibody diversity is generated in the B lymphocytes of the immune system using recombination, gene rearrangement and mutation. Immature B cells found in bone marrow assemble immunoglobulin (Ig) genes in a process called V(D)J recombination catalysed by the complex of rearrangement activating genes RAG1-RAG2. The exon encoding the antigen-binding domain is assembled from V (variable), D (diversity) and J (joining) gene segments and this V(D)J segment is linked to the μ constant region (Cμ). This generates a primary repertoire of the low affinity IgM antibodies.

After antigen contact, B cells expressing an IgM molecule that recognizes the antigen are induced to proliferate and form germinal centres in secondary lymphoid organs. The antigen–binding regions undergo further diversification by somatic hypermutation (SHM). This introduces point mutations along the variable region allowing generation of antibodies with improved affinity for the antigen. In addition to this, stretches of nucleotide sequences are copied from the pseudogene V elements (ψ V) into the functional V(D)J exon further modifying the variable region in a process called immunoglobulin gene conversion (IGC). Finally, exons encoding the constant region of the antibody are swapped by recombination events replacing the Cμ constant region of the IgM antibody with the γ, ε or α constant regions allowing the production of IgG, IgE and IgA antibodies. This final process is known as class switch recombination (CSR).

1.2.2.2 Activation-induced deaminase

Activation-induced deaminase (AID) was discovered by Muramatsu *et al.* using subtractive hybridisation to identify genes in murine B lymphoma cells whose expression increased in parallel with the induction of class switch recombination (CSR) (Muramatsu et al., 1999).

Basal expression of AID is low however, upon stimulation by IL-4, TGF-β and CD40L, 80% of B cells switch from IgM to IgA and AID expression is up regulated by 10 fold. This increase in AID mRNA is synchronous with germinal centre formation after antigen stimulation. Cyclohexamide treatment reduced this up regulation of AID after stimulation indicating *de novo* synthesis of proteins is also required (Doi et al., 2003; Muramatsu et al., 1999).

One open reading frame was identified encoding a 198 residue protein with an approximate molecular weight of 24 kDa. The murine AID protein shows 34% identity to APOBEC-1 and contains the well characterised cytidine deaminase (CDA) motifs and a leucine-rich region at the carboxy terminus. Like APOBEC-1, AID has also been shown to shuttle between the nucleus and cytoplasm via a N-terminal nuclear localization signal and a C-terminal nuclear export signal (Ito et al., 2004).

AID has been demonstrated to have CDA activity on [^{3}H] deoxycytidine. Tetrahydrouridine, a CDA specific inhibitor, and 10-o-phenanthroline, a zinc chelator, inhibit this activity indicating that the reaction is zinc dependent (Muramatsu et al., 1999).

The human *AID* gene has been mapped to chromosome 12p13, very close to the APOBEC-1 locus (12p13.1 – p13.2) indicating that the two genes may originate from a common ancestor that has CDA activity. A 2.9 kb AID mRNA is expressed in B lymphocytes but not seen in fibroblasts or T cells.

Human AID shows 92% amino acid identity to the mouse protein, especially at the CDA motifs and C-terminal leucine-rich region (Muto et al., 2000).

1.2.2.3 AID deletions

AID $^{-/-}$ mice and AID $^{-/-}$ / RAG-2 $^{-/-}$ chimera were generated by Muramatsu *et al.*. Phenotypically both mice were similar. B, T and other haematopoietic cells all developed normally. At 10 weeks neither serum IgG3, IgG2b nor IgA were detected and low levels of IgG1 and IgG2a were observed but found to be maternal in origin. Interestingly, serum IgM levels were found to be 2-3 fold higher in AID $^{-/-}$ mice than AID $^{+/-}$ mice (Muramatsu et al., 2000).

When exposed to T-dependent antigens, antigen-specific IgM was produced but no IgG1 antibody was seen. This indicates that AID deficiency causes a complete defect in class switching and the mice show a hyper-IgM phenotype with enlarged germinal centres containing strongly activated B cells before and after immunization. *In vitro* studies in AID $^{-/-}$ B cells confirm the phenotype is not due to a defect in B cell activation but that AID is required for CSR and overexpression of AID in B lymphoma cells results in an increase of class switching from IgM to IgA without cytokine stimulation. In addition to this, the VH gene did not accumulate mutations after immunization signifying a defect in somatic hypermutation (Muramatsu et al., 2000; Nagaoka et al., 2002). Studies on AID $^{-/-}$ chicken DT40 cells also indicate a complete block of Ig gene conversion (Arakawa et al., 2002).

The AID $^{-/-}$ mouse phenotype of no class switch recombination, no somatic hypermutation and lymph node hyperplasia resembles that of the autosomal recessive Hyper-IgM Syndrome (HIGM2). Linkage analysis maps the HIGM2 locus to chromosome 12p13, the same locus as human *AID*. The *AID* gene in a number of HIGM2 patients was sequenced and 10 independent mutations described. Some mutations are predicted to lead to truncated proteins, others are missense mutations localised within the putative cytidine

deaminase catalytic region (Revy et al., 2000; Zhu et al., 2003). These studies suggest class switching, somatic hypermutation and Ig gene conversion all share a common molecular mechanism involving the AID protein.

1.2.2.4 Mechanism

The mechanism by which AID initiates CSR, SHM and IGC was initially thought to be through the editing of mRNAs. This hypothesis was given credence by the striking homology between AID and the RNA editing enzyme APOBEC-1 and that AID requires *de novo* protein synthesis to initiate CSR (Doi et al., 2003; Muramatsu et al., 1999). However, AID showed no CDA or binding activity on the APOBEC-1 substrate apoB and was unable to bind to an AU-rich RNA (Muramatsu et al., 1999).

An alternative hypothesis proposed that AID acted on the DNA at the immunoglobulin gene locus either directly by deamination or indirectly through one or more auxiliary factors. Indeed, when expressed in *E. coli* AID was shown to deaminate dC residues (Petersen-Mahrt et al., 2002) in single strand DNA (Bransteitter et al., 2003; Chaudhuri et al., 2003; Dickerson et al., 2003). Futher studies suggest a link between AID-mediated hypermutation and transcription with transcription targeting AID to dsDNA by generating secondary structures that provide the necessary ssDNA substrates (Bransteitter et al., 2003; Chaudhuri et al., 2003).

The model proposed suggests that the recruitment of AID to the Ig locus results in the localised deamination of dC – to – dU. Repair of this uracil can occur by a number of methods and the pathway used and location of the lesion determines whether this leads to SHM, IGC or CSR (Petersen-Mahrt et al., 2002).

1.3 A-I EDITING

1.3.1 ADAR 1

The ADAR1 enzyme was initially purified from bovine liver (Kim et al., 1994a) and calf thymus (O'Connell and Keller, 1994). The purified proteins exhibited adenosine–to–inosine editing activity with the 116 kDa polypeptide (O'Connell and Keller, 1994) showing twice the specific activity of the smaller 93, 88 and 83 kDa polypeptides (Kim et al., 1994a). The protein was sensitive to N-ethymaleimide (NEM) and required 1-10 mM DTT for activity indicating the enzyme may contain essential SH group(s) perhaps involved in the catalytic mechanism.

Human ADAR1 cDNA was isolated by Kim *et al.* and O'Connell *et al.* from a HeLa cDNA library. The full-length cDNA clone was found to give a single open reading frame of 1226 amino acids with a predicted molecular weight of 139 kDa (Kim et al., 1994b; O'Connell et al., 1995). The human *ADAR1* gene has been localised to chromosome 1q21 and contains 15 exons spanning 30kb (Wang et al., 1995). It appears to be ubiquitously expressed in human tissue (Wagner et al., 1990) with relatively high levels in the brain (Kim et al., 1994b).

Analysis of the primary structure indicates the presence of three dsRNA binding domains and a C-terminal region containing the HAE and PCG motifs commonly found in the cytidine deaminase family (Kim et al., 1994b; Liu and Samuel, 1996; O'Connell et al., 1995) (Figure 1.4). Further analysis identified two putative Z-DNA binding motifs at the amino terminus (Herbert et al., 1997). The Zα motif has been shown to bind Z-DNA in a conformation-specific manner (Herbert et al., 1998) whilst the Zβ does not bind Z-DNA on its own, only when part of the larger Zab domain. This 229 amino acid region

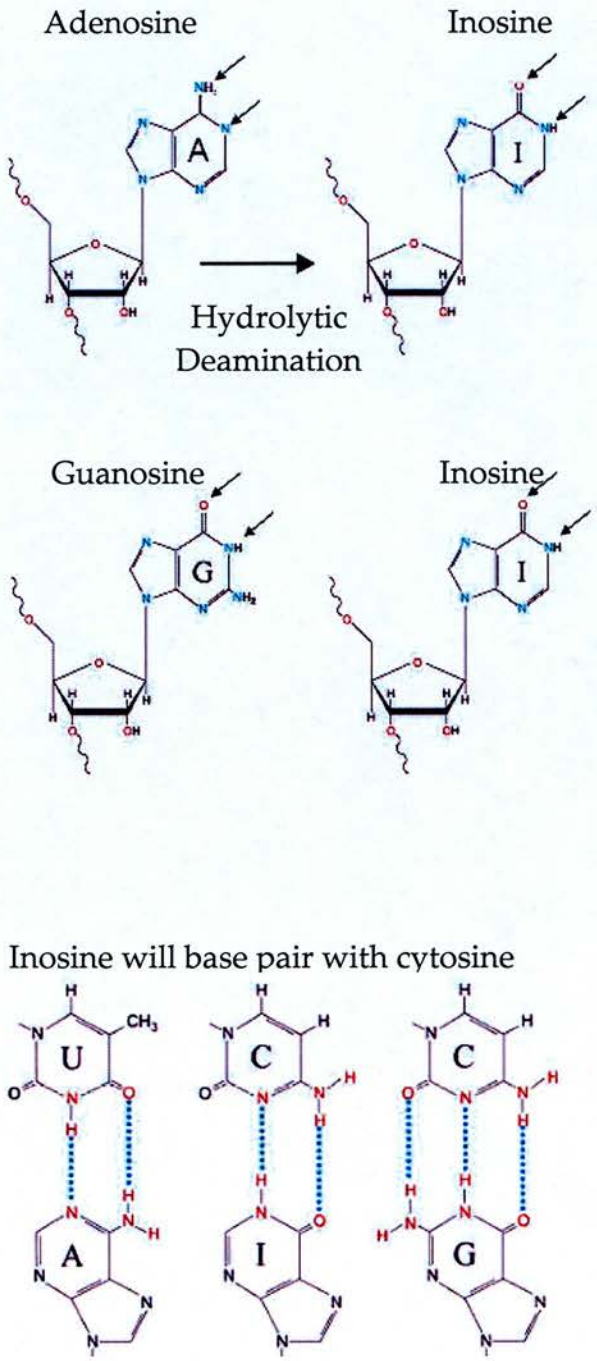


Figure 1.3 Adenosine deamination by hydrolytic deamination.
ADARs catalyse a hydrolytic deamination at C6 of adenosine to inosine. Inosine structurally resembles guanosine, will form base pairs with cytosine and is read as guanosine by the translational machinery.

consists of Zα and Zβ motifs separated by a tandem repeat of a 49 amino acid linker (Schwartz et al., 1999).

It has been postulated that Z-DNA formation occurs 5′ to or behind a moving RNA polymerase during transcription so recognition of Z-DNA by ADAR1 provides a plausible mechanism by which ADAR1 can be targeted to nascent RNA so that editing occurs before splicing however, this method of regulating enzyme activity has yet to be demonstrated.

A screen to isolate interferon-regulated cDNAs identified a 6.7 kb clone designated K88, whose expression was increased fivefold by interferon treatment (Liu and Samuel, 1996; Patterson et al., 1995). Immunoblot analysis using antisera prepared against K88 cDNA products recognized an interferon-inducible 150 kDa protein associated with both the nucleus and cytoplasm and a second constitutively expressed primarily nuclear 110 kDa protein. Western blot analysis and adenosine deaminase activity assays illustrated that both the p150 and p110 proteins were dsRNA-specific adenosine deaminases. However, the full-length p150, using the translation start site M1, was poorly expressed in COS cells when compared to the p110 isoform using the M296 translation start site (Patterson and Samuel, 1995). Three promoters have been identified upstream of the *ADAR1* gene (George and Samuel, 1999; George et al., 2005). The interferon inducible promoter, P_I, has been shown to contain elements known to be interferon responsive. This promoter region generates transcripts incorporating exon 1A that contains the M1 translation start site. This results in a single open reading frame of 1226 amino acids (iADAR1). The constitutively active promoter, P_C, has no interferon responsive elements and is immediately 5′ to exon 1B. Transcripts initiated from this promoter incorporate exon 1B structures giving an open reading frame of 931 amino acids initiating from the alternative start site M296 found in exon 2 (cADAR1). The amino terminal truncated M296 ADAR1 protein contains all three dsRNA binding motifs and the catalytic

deaminase domain. However, it lacks the Zα motif of the Zab domain though the biological significance of this is unknown (George and Samuel, 1999). A third promoter has recently been identified. Incorporating exon 1C, this transcript is very scarce and only detected in the brain and heart (George et al., 2005).

cDNA library screens with ADAR1 cDNA have also identified two alternative splice sites involving a 78 nucleotide deletion in exon 7 and a 57 nucleotide deletion in exon 6. These splicing events alter the spacing between the third dsRNA binding motif and the catalytic domain and the second and third dsRNA binding motifs, respectively. Three ADAR spliceoforms were found to be differentially expressed in human cells. ADAR-a containing full-length exons 6 and 7, ADAR-b containing full-length exon 6 with the exon 7 deletion and ADAR-c containing both deletions. The fourth possible variant consisting of full-length exon 7 and the exon 6 deletion was not detected in any cDNA library examined but was expressed *in vivo* and *in vitro* alongside the three naturally occurring variants in both the full-length M1 and the shorter M296 forms. All recombinant variants possessed comparable dsRNA deaminase activity indicating that the deletions resulting from alternative splicing in exons 6 and 7 have no discernible effect on enzyme activity *in vitro* (Liu et al., 1997).

Deletion analysis and point mutations of the ADAR1 enzyme revealed that the first and third dsRBMs are important for dsRNA binding with the third dsRNA binding domain being essential (Lai et al., 1995; Liu and Samuel, 1996). In spliceoforms with the exon 7 deletion the second dsRNA binding motif was found to be dispensable whilst the presence of both the exon 6 and exon 7 deletions appears to reduce the functional significance of the first dsRNA binding motif (Liu et al., 1997). Point mutations within the catalytic domain demonstrated that the histidine (H910) and cysteine (C966) residues found within the deaminase CHAE and PCG motifs and another cysteine (C1036) are essential for activity and probably play a role in the catalytic

mechanism. A glutamate residue at position 912 was predicted, based on the three-dimensional structure of *E. coli* cytidine deaminase, to be involved in the critical proton transfer function of the enzyme. A substitution at this position with alanine also abolishes all activity (Lai et al., 1995; Liu and Samuel, 1996). Any deletion at the carboxy terminus has been shown to abolish enzyme activity (Lai et al., 1995).

A study by Cho *et al.* demonstrated that ADAR1 exists predominantly as a homodimer. Like the *E. coli* cytidine deaminase and APOBEC-1, the ADAR1 monomers were found to interact over a widespread region of the protein including the deaminase domain and dsRNA binding domain with the N-terminus being dispensable (Cho et al., 2003). This region of interaction differs greatly from that of other dsRNA binding domain-containing proteins such as PKR and *Drosophila* ADAR which mediate their dimerisation via N-terminal regions including their dsRNA binding domains (See section 1.6.2). ADAR1 did not form heterodimers with either ADAR2 or ADAR3 (See sections 1.3.3 and 1.3.5) even though previous studies showing the mixing of ADAR3 with either ADAR1 or ADAR2 resulting in diminished editing activity suggested this may be a possibility (Chen et al., 2000).

The nuclear localisation signal of the human ADAR1 is not located in the N-terminus as previously believed but overlaps with the third dsRNA binding motif. This NLS activity is independent of RNA binding indicating a dual function for this region of protein. Deletion analysis demonstrated that only the entire dsRNA-binding motif could exhibit NLS activity (Eckmann et al., 2001). The nuclear localisation of the human protein appears to be regulated in a transcription-dependent manner similar to that of the hnRNPs. Cells treated with actinomycin D showed a reduced concentration of hADAR1 in the nucleus and accumulation in the cytoplasm indicating that the protein

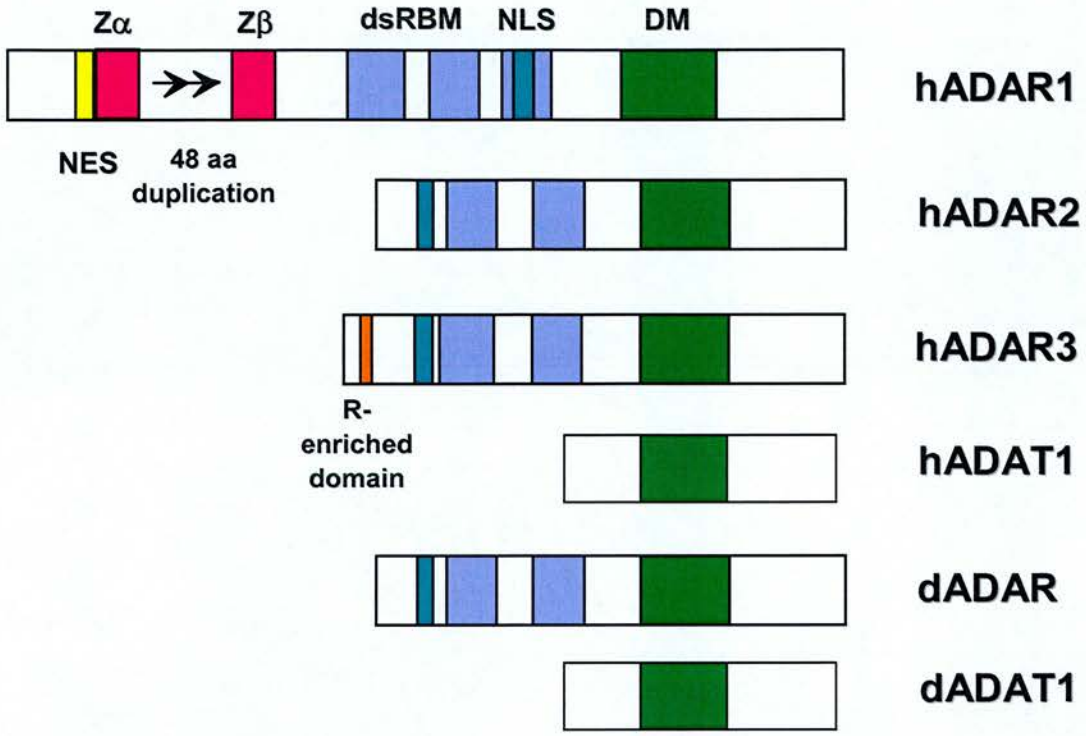


Figure 1.4 Domain structure of the Adenosine deaminase family of proteins. NES, nuclear export signal; Zα and Zβ, Z-DNA binding domains; dsRBM, dsRNA binding domain; NLS, nuclear localisation signal; DM, deaminase domain.

shuttles between the two compartments and that nuclear localisation is not solely dependent on the presence of a NLS but requires on-going RNA polymerase II transcription (Eckmann et al., 2001).

The nuclear export signal (NES) of iADAR1 is located in a leucine rich region of the proteins N-terminus and mutations in these key hydrophobic residues results in nuclear localisation of ADAR1 (Poulsen et al., 2001). The nuclear export of iADAR1 is inhibited by leptomycin B, a drug that covalently binds the nuclear export receptor CRM-1, impeding its function. This suggests that iADAR1 utilises a CRM1-dependent mechanism for nuclear export and this is further supported by the fact that CRM1 specifically binds to the NES of iADAR1 (Poulsen et al., 2001). It should be noted that this NES is not present in the cADAR1 isoform perhaps accounting for it nuclear localisation (Poulsen et al., 2001).

ADAR1 has also been shown to shuttle between the nucleolus and nucleoplasm with this transient nucleolar sequestration thought to be mediated by the abundance of dsRNA structures within this compartment (Desterro et al., 2003). iADAR1 is located between the cytoplasm and nucleus with cADAR1 locating to the nucleus and nucleolus due to its lack of a nuclear export signal (Poulsen et al., 2001). Rather than editing substrates being recruited to the nucleolus, expression of an editing substrate was found to delocalise ADAR1 suggesting this nucleolar localisation is transient and dynamic (Desterro et al., 2003).

Whilst in the nucleolus, ADAR2 (See section 1.3.3) has been shown to edit transcripts whereas ADAR1 appears inactive (Vitali et al., 2005). Desterro *et al.* recently reported that ADAR1 co-localises with the small ubiquitin modifier, SUMO-1, in a specific sub-nucleolar region (Desterro et al., 2005). Analysis of the ADAR1 peptide sequence indicated the presence of 2 potential SUMO-1 consensus motifs at lysines 418 and 899. ADAR1 was shown to be modified by SUMO-1 at lysine 418 with a K418R point mutation

abolishing this sumoylation. This modification is not a requirement for ADAR1 function but instead has been shown to reduce the editing activity of the enzyme indicating that sumoylation may play a role in ADAR1 enzyme regulation (Desterro et al., 2005). As ADAR2 is not modified by SUMO-1, the modification and resulting decrease in activity of ADAR1 could explain the absence of ADAR1-mediated editing within the nucleolus.

Many groups have examined the role of the interferon-inducible promoter controlling ADAR1 expression. Systemic inflammation has been shown to increase the number of inosine-containing mRNAs, a direct consequence of increased ADAR1 expression and subsequent increase in A-to-I editing (Yang et al., 2003). Studies on lung tissue have demonstrated that ADAR1 expression is upregulated when microvascular injury is induced with pulmonary expression of ADAR1 localising to inflammatory cells including neutrophils and monocytes (Rabinovici et al., 2001). Interferon is produced early on in the inflammation cascade suggesting that the M1 form of ADAR1 plays a role in the inflammatory response perhaps through the modulation of protein production.

A genome wide search in three families suffering from the autosomal dominant disorder Dyschromatosis symmetrica hereditaria (DSH) mapped the disease locus to chromosome 1q21.3 and identified 4 mutations within the *ADAR1* gene in affected families (Miyamura et al., 2003); two nonsense mutations predicted to result in truncated proteins, a leucine to proline amino acid change in the deaminase domain that probably compromises enzymatic activity and, a phenylalanine to serine change in exon 15 at the carboxy terminus. A further seven novel mutations have since been identified (Zhang et al., 2004). This disorder is characterised by a mixture of hyperpigmented and hypopigmented macules distributed on the dorsal aspects of the hands and feet, a phenotype not seen in any *adar*$^{+/-}$ or *adar*$^{-/-}$ mouse models that have been generated. In fact the distinct regions of

altered pigmentation are difficult to explain pathologically mainly due to the ubiquitous expression of *ADAR1*.

A recent study has suggested an involvement with the RNAi pathway (Nie et al., 2004). The long form of ADAR1 is found in the cytoplasm and has been show to bind very specifically and strongly to siRNA in a sequence-independent manner. ADAR2 and the shorter form of ADAR1 have also been shown to bind to siRNA but with much lower affinities. All the ADARs bound to siRNA molecules as small as 15 bp but A-to-I editing only occurred in molecules ≥30 bp. The efficiency of RNAi silencing was greatly enhanced in mouse embryonic fibroblasts null for *ADAR1* or expressing a mutant ADAR1 protein not containing the dsRNA binding domains. These results suggest that the inhibitory effect ADAR1 has on the RNAi pathway is mediated via its dsRNA binding domains and its ability to sequester rather than edit siRNA (Zhang et al., 2004). Indeed, a subunit of the RNA-induced silencing complex (RISC), Tudor staphylococcal nuclease (Tudor-SN), has been shown to interact with hyper-edited dsRNA substrates promoting their cleavage (Scadden, 2005). This is the first evidence of a direct interaction between both pathways.

1.3.2 ADAR 1 deficient mice

Initial attempts by Wang *et al.* to generate *adar1*$^{-/-}$ transgenic mice by targeting exons 12 and 13, corresponding to a part of the catalytic domain, proved impossible. *adar1*$^{+/-}$ chimeric mice did not survive past embryonic day 14 and displayed defects in the haematopoietic system suggesting the presence of substrates in the liver that are essential for normal development (Wang et al., 2000). The likely existence of a dominant-negative truncated protein inhibiting already reduced levels of wild-type enzyme could explain the heterozygous-lethal phenotype observed.

In order to minimize the potential for interfering products, new null alleles were constructed in which exons 2-13 (Hartner et al., 2004) were replaced with a *pgk-neo* gene. These exons encode residues corresponding to all three dsRNA binding domains, the two putative Z DNA binding domains and most of the catalytic deaminase domain. *adar1*$^{+/\Delta 2\text{-}13}$ heterozygous mice appeared normal and intercrosses produced heterozygous and wild-type pups at normal frequencies but no homozygous pups indicating that differently constructed *adar1* null alleles result in embryonic death between E11.5 and E12.5 in homozygotes. Interestingly, Hartner *et al.* generated *adar1*$^{+/\Delta 7\text{-}9}$ mice that appeared normal and went on to produce *adar1*$^{\Delta 7\text{-}9/\Delta 7\text{-}9}$ embryos that phenotypically resembled the *adar1*$^{\Delta 2\text{-}13/\Delta 2\text{-}13}$ mice. Indeed, Wang *et al.* (2004) have recently successfully generated both *adar1*$^{+/-}$ and *adar1*$^{-/-}$ mice with exons 12-15 excised via the Cre-LoxP system. Therefore the haploinsufficiency seen in Wang *et al.* (2000) seems to have resulted from insufficient expression of the manipulated *adar1* allele rather than a dominant-negative truncated protein.

To determine the effect of ADAR1 deficiency on RNA editing, a number of editing sites were analysed. No editing of sites A and B of the 5-HT$_{2C}$ serotonin receptor was detected in *adar1*$^{-/-}$ mice compared with levels of 80-90% in wild-type mice (Hartner et al., 2004; Wang et al., 2004). There was also a reduction in editing of GluR5 Q/R and GluR-B R/G indicating that they too may be potential targets of ADAR1 (Wang et al., 2004). Editing at site C of the 5-HT$_{2C}$ receptor and Q/R site of GluR-B remained unchanged whilst sites D and E of the 5-HT$_{2C}$ receptor showed increased levels of editing when compared to wild-type (Hartner et al., 2004; Wang et al., 2004). These results may reflect easier access for ADAR2 to these sites in the absence of ADAR1 and suggests a site-selective A-to-I editing function for ADAR1 indicating that the ADAR proteins can edit distinct subsets of adenosines even when they are closely grouped together.

ADAR1-deficient mice die *in utero* between E11.0 and E12.5 (Hartner et al., 2004; Wang et al., 2004). They appear grossly normal except for a significant reduction in foetal liver size and a pale yolk sac and embryo when compared to wild-type. The vascular architecture of yolk sacs appears normal with vessels containing nucleated erythrocytes with no morphological abnormalities.

The first wave of haematopoiesis is extra-embryonic and takes place in the yolk sac from around E7 to E11 producing nucleated erythrocytes termed primitive. At around E11, the foetal liver becomes the major haematopoietic tissue and gives rise to blood cell lineages (termed definitive) found in the adult until shortly before birth when haematopoiesis shifts to the spleen and bone marrow. At E11.5 of *adar1*$^{-}$/$^{-}$ embryos cell numbers in embryonic haematopoietic tissues are significantly reduced in the yolk sac and peripheral blood when compared to *adar1*$^{+}$/$^{-}$ and wild-type embryos. Analysis of E11.5 *adar1*$^{-}$/$^{-}$ embryos' yolk sacs, foetal livers and peripheral blood contain significantly fewer definitive (foetal liver-derived) haematopoietic progenitor cells when compared with *adar1*$^{+}$/$^{-}$ and wild-type embryos (Hartner et al., 2004). In contrast to this, Wang *et al.* (2004) found no significant difference in the number of definitive erythroid and myeloid progenitor cells in the yolk sacs of *adar1*$^{-}$/$^{-}$, *adar1*$^{+}$/$^{-}$ and wild-type embryos at E9.5 (Wang et al., 2004). This also contradicts the near complete block of definitive erythropoiesis they observed previously (Wang et al., 2000).

Rather than the dramatic increase in erythroid cells seen in wild-type foetal livers at E11.25, ADAR1-deficient livers contain very few erythroid cells, most of which are yolk sac derived nucleated erythrocytes (Hartner et al., 2004). ADAR1-deficient livers show a profound scarcity of haematopoietic cells at E11.25 onwards, with many cells displaying the characteristic signs of apoptotic cell death (Hartner et al., 2004). TUNEL assays revealed that these elevated levels of apoptosis can be seen in many tissues including vertebra and heart and have been detected as early as E10.5. The highest levels of

apoptosis were detected within the foetal liver for both hepatocytes as well as the haematopoietic cell population (Wang et al., 2004). This may explain the difference in numbers of definitive progenitor cells seen on days E9.5 and E11.5.

From E11.5 onwards, there is a significant reduction in cell density of the liver with blood accumulating in pathologically enlarged spaces perhaps created by cell death. A reduction in cell proliferation to about 30% of wild-type as revealed by bromodeoxyuridine incorporation (Hartner et al., 2004) accompanied by high levels of apoptosis in the hepatocyte population (Wang et al., 2004) contribute to this breakdown of liver architecture. No morphological abnormalities were found in other tissues including neuroepithelium, heart, branchial arches, placenta and yolk sac (Hartner et al., 2004).

The mechanism by which ADAR1 prevents liver disintegration and increased apoptosis in wild-type embryos is unclear. Studies in mouse embryonic fibroblasts indicate that apoptosis seen in *adar1* null embryos is stress-induced and similar conditions result in the significantly increased expression of the ADAR1 p150 isoform in wild-type MEF cells. Apoptosis induced in *adar1* null embryos is independent of the PKR and RNase L pathways both of which are activated by the presence of dsRNA and can lead to the induction of apoptosis in cells under stress (Wang et al., 2004). No haematopoietic or liver transcripts have been identified that need ADAR1-mediated A-to-I editing therefore this anti-apoptotic ADAR1 function may involve a new and novel mechanism.

1.3.3 ADAR 2

In an attempt to find the dsRNA adenosine deaminase activity responsible for GluR-B Q/R editing, a screen of rat brain cDNA using a probe based on the predicted catalytic domain of ADAR1 was performed. A number of

clones were identified, one of which was capable of editing both the Q/R site and R/G sites of GluR-B as well as converting adenosine to inosine on synthetic dsRNA molecules (Melcher et al., 1996b). Initially named RED1, this clone shared a 31% overall sequence identity to ADAR1 and consisted of 2 dsRNA binding motifs and a C-terminal catalytic domain where the cytidine deaminase motifs, CHAE and PCG were present (Melcher et al., 1996b) (Figure 1.4).

A screen of human libraries isolated a number of different cDNA clones, each of which was derived from alternative splicing of the same original transcript (Gerber et al., 1997; Lai et al., 1997) and mapped to the long arm of chromosome 21, region q22.2-q22.3 (Cheng et al., 1994). All isoforms share an identical N-terminus, two dsRNA binding motifs and a catalytic domain . However, two isoforms differ by the presence of an *Alu* cassette in the middle of the deaminase domain (Gerber et al., 1997; Lai et al., 1997). The identification of putative exonic splice sites in this region suggest the *Alu* cassette is alternatively spliced and *in vitro* assays demonstrate that the short form, the isoform minus the *Alu* cassette, has a higher catalytic activity than the long form (Gerber et al., 1997).

A number of other clones differed at the 3'-end with some failing to encode the C-terminal 29 amino acids (Gerber et al., 1997; Lai et al., 1997) and these isoforms were found to display no editing activity (Lai et al., 1997) indicating residues in this region may be involved with protein stability and/or enzymatic function.

Rat ADAR2 was found to modify its own transcript creating a new AG 3' splice site (Rueter et al., 1999). This alternative splicing results in the inclusion of an additional 47 nucleotide region, a frameshift and the predicted translation of an 88 amino acid polypeptide (Rueter et al., 1999). In rat brain a high percentage of mRNAs containing this 47 nucleotide region were detected however, low levels of the truncated protein are seen

suggesting that much of the edited mRNA is targeted for nonsense mediated decay (Rueter et al., 1999).

hADAR2, the human 90 kDa protein was originally purified from HeLa nuclear extract (O'Connell et al., 1997). The enzyme was sensitive to *N*-ethyl-maleimide (NEM) and its activity inhibited by *o*-phenanthroline by 30% when compared to the 90% inhibition seen with ADAR1. This suggests the presence of a Zn^{2+} ion at the active site of both enzymes which is less accessible in ADAR2 than in ADAR1 (O'Connell et al., 1997).

Kinetic analysis of the ADAR2 editing reaction indicates that the protein functions as a dimer (Jaikaran et al., 2002). Further studies confirmed this result stating that like ADAR1, ADAR2 exists predominantly as a homodimer with the monomers interfacing over both the catalytic and dsRNA binding domains (Cho et al., 2003) (See section 1.6.2).

ADAR2 was found to accumulate in the nucleolus (Desterro et al., 2003; Sansam et al., 2003) and required binding to dsRNA for this localisation (Sansam et al., 2003). FRAP analysis indicates that ADAR2 shuttles between the nucleolus and nucleoplasm with actinomycin B treatment resulting in a complete translocation of ADAR2 to the nucleoplasm (Desterro et al., 2003; Sansam et al., 2003). Deletion analysis indicated the presence of a nucleolus targeting signal in the N-terminal region of ADAR2 involving residues 75 to 132 that is necessary but not sufficient for nucleolar targeting (Desterro et al., 2003).

These results suggest the nucleolus is a site of functional sequestration for ADAR2 activity and limits its nucleoplasmic activity on endogenous pre-mRNAs. Expression of an editing substrate was found to delocalise ADAR2 from the nucleolus suggesting the localisation of ADAR2 to this compartment is transient (Desterro et al., 2003). More recent studies however, have shown that there is no ADAR1-mediated editing observed

within the nucleolus but ADAR2 can edit sites in the GluR-B and 5-HT_{2c} substrates (Vitali et al., 2005). One of these targets is the C-site of the serotonin receptor and nucleolar editing at this site has been shown to be specifically inhibited by a C/D small nucleolar RNA, MBII-52. This snoRNA is believed to target the C-site for 2'-*O*-methylation and thus, as discussed by Yi-Brunnozzi *et al.*, decrease the efficiency of editing at this site (Yi-Brunozzi et al., 1999).

An as yet unexplained reduction in ADAR2-mediated editing of GluR-B has been implicated in the selective neuronal death seen in sporadic amyotropic lateral sclerosis or ALS (Kwak and Kawahara, 2005). GluR-B editing remains normal in other cell types but is markedly reduced to approximately 50% within the motor neurons of patients with ALS indicating that the resulting increase in Ca^{2+} permeability may play a role in the selective motor neuron death seen in these patients (Kawahara et al., 2004; Kwak and Kawahara, 2005).

1.3.4 ADAR 2 deficient mice

adar2^{+}/$^{-}$ and *adar2*$^{-}$/$^{-}$ transgenic mice were generated by the insertion of a *pgk-neo* gene in exon 4 that contains the essential deaminase motif (Higuchi et al., 2000). The *adar2*$^{+}$/$^{-}$ mice appeared normal however the homozygotes suffered seizures and died between 0 and 20 days after birth. Electrophysiological analysis showed an increase in AMPA receptor conductance and this appears to be the primary cause of seizures.

Analysis of 25 known editing sites indicated that editing was substantially reduced at most positions in *adar2*$^{-}$/$^{-}$ mice, in particular the GluR-B Q/R site where editing was reduced 10-fold. The GluR-B primary transcripts were incompletely processed resulting in nuclear accumulation and a 5-fold reduction in GluR-B mRNA. This suggests that editing is a prerequisite for efficient splicing and processing of the GluR-B pre-mRNA.

GluR-B$^{+/\Delta ECS}$ transgenic mice were generated by the deletion of the downstream editing site complementary sequence (ECS) in a GluR-B allele preventing editing at the Q/R site (Brusa et al., 1995). The absence of editing at the Q/R site resulted in a shortage of edited GluR-B subunits for assembly of the heteromeric AMPA receptor, an increase in glutamate-activated Ca^{2+} permeability and death (Brusa et al., 1995).

The *adar2*$^{-/-}$ phenotype was rescued by introducing a *GluR-B*^{R} codon containing an arginine residue at the Q/R site. Heterozygous *adar2*$^{-/-}$*/GluR-B*$^{+/R}$ mice showed partial rescue with the phenotype less impaired when compared to *GluR-B*$^{+/\Delta ECS}$ mice (Brusa et al., 1995). This is consistent with the theory that phenotypic impairment increases with the Q/R ratio of GluR-B. As one GluR-B allele expresses exclusively the "R" form, the primary transcripts from the second allele are under edited and mRNA levels are reduced due to impaired splicing and processing. *adar2*$^{-/-}$*/GluR-B*$^{R/R}$ mice appear normal at all ages (Higuchi et al., 2000). These results identify the GluR-B Q/R site as the most important target of ADAR2.

It should be noted that GluR-B deficient mice are mostly viable with the phenotype less severe than that of ADAR2 deficient mice even though the Ca^{2+} permeability of the AMPA receptors is greatly increased. This difference may be due to the possibility that the GluR-B subunit may enhance AMPA receptor assembly (Feldmeyer et al., 1999). Recent studies have demonstrated that this is indeed the case and Q/R editing modulates the assembly of AMPA receptor tetramers with the GluR2 (R) isoform being retained in the endoplasmic reticulum. Only when assembly is possible with other subunits does GluR2 (R) exit the ER and proceed to the cell surface thus precluding the formation of homomeric GluR2 (R) channels (Greger et al., 2003).

1.3.5 ADAR 3

ADAR3 cDNA was first isolated from rat brain total RNA. The full-length cDNA clone was found to encode a 746 amino acid protein with a predicted molecular weight of 82 kDa. The human gene was mapped to chromosome 10p15 (Mittaz et al., 1997) and the human protein was shown to be 72% similar in sequence to human ADAR2 having a similar domain structure with 2 dsRNA binding motifs and a catalytic deaminase domain (Chen et al., 2000) (Figure 1.4). However, unlike ADAR2, ADAR3 contains an N-terminal 54 amino acid region rich in arginine and lysine residues known as the R domain (Melcher et al., 1996a). ADAR3 transcripts appear to be restricted to the brain in both humans and rats, with the thalamus and amygdala containing relatively higher levels (Chen et al., 2000).

Melcher *et al.* and Chen *et al.* demonstrated that the purified rat and recombinant human ADAR3 enzyme have no activity on well characterised ADAR substrates such as the GluR-B "Q/R" site, "R/G" site, and the intronic "hot spot" site as well as synthetic dsRNA (Chen et al., 2000; Melcher et al., 1996a). However, filter binding assays show ADAR3 is capable of binding both dsRNA and ssRNA, with the former mediated by the 2 dsRNA binding motifs and the latter being mediated by the R domain (Chen et al., 2000). This RNA binding activity may account for the inhibitory effect ADAR3 has on ADAR1 and ADAR2 editing activity *in vitro* (Chen et al., 2000).

ADAR3 may exhibit no activity *in vitro* but the fact that it can bind RNA and has conserved motifs within the catalytic domain, suggests that this is a functioning enzyme with possible substrates waiting to be discovered. The likelihood is that ADAR3 acts as a competitive inhibitor suppressing the activity of the other ADAR enzymes either by competitive substrate binding or by sequestering the active ADARs in non-functioning heterodimers (Chen et al., 2000; Cho et al., 2003).

1.3.6 *Drosophila* ADAR

dsRNA adenosine deaminase activity was first identified in *Drosophila* extracts by Casey *et al.* (Casey and Gerin, 1995) and a *Drosophila* cDNA clone with homologous to ADAR1 and ADAR2 was first isolated using degenerate primers based on the highly conserved ADAR catalytic domain (Hanrahan et al., 1998).

Palladino *et al.* also identified a PCR product from a *Drosophila* whole adult cDNA library with significant homology to the ADAR deaminase domain and higher homology to ADAR2 than ADAR1. This was identified as the *Drosophila Adar* gene and localised to the 2B6-7 region of the X chromosome (Palladino et al., 2000a).

Sequence analysis of cDNA clones identified a number of start methionines, the exon -1 and +1 ATG sites being the most frequently used (Palladino et al., 2000a). Two promoters were identified generating either exon -4a or -4b-containing transcripts. The -4a promoter was active throughout all stages of fly development unlike the -4b promoter which appears to be highly regulated (Palladino et al., 2000a). Alternative splicing was observed at exon 3a between the two dsRNA binding motifs and alters the spacing between these structural motifs (Palladino et al., 2000a).

Overall, 4 major transcripts were identified, generated by the presence or absence of exons -1a and 3a with the exclusion of all other alternative exons (Palladino et al., 2000a). cDNA sequence analysis predicted a protein of 670 amino acids and on purification, an 80 kDa band was observed that exhibited A-to-I editing activity on synthetic dsRNA substrates (Palladino et al., 2000a). Indeed, four dADAR protein isoforms can be translated starting at the exon –1 or exon +1 ATG and including or excluding 3a.

The dADAR enzyme was found to edit its own transcript. This editing event occurs at a site in exon 7 within the catalytic domain six amino acids

downstream of the second cytidine deaminase motif (PCGD) and results in a serine to glycine amino acid change (Palladino et al., 2000a). Termed the S/G site, editing here, as with other *Drosophila* ADAR substrates (Hanrahan et al., 2000; Keegan et al., 2005), is highly regulated with low levels seen in the *Drosophila* embryo and pupae (Palladino et al., 2000a). Self-editing increases dramatically upon eclosion with over a 40-fold increase from embryo to adult. This serine to glycine amino acid change is essential for *Drosophila* viability and results in a dADAR isoform less active than the genomically encoded serine isoform (Keegan et al., 2005). Interestingly, if editing at this site is prevented *in vivo* and only genome-encoded ADAR is present, hyperediting occurs at other physiological targets leading to early death in the flies (Keegan et al., 2005).

Recent studies have demonstrated that like ADAR1 and 2, dADAR functions as a dimer (Gallo et al., 2003). Unlike ADAR1, the deaminase domain is not required, instead in a pattern resembling other dsRBD-containing proteins, the region of interaction was identified as the N-terminus and first dsRBM (Gallo et al., 2003). Abolition of RNA binding activity in dADAR by point mutations within the first dsRBD disrupted dimerisation indicating the requirement of dsRNA for dimer formation (Gallo et al., 2003). Various dADAR isoforms were able to interact and generate heterodimers exhibiting different editing activities suggesting that alongside the alternative splicing of exon 3a and self-editing, dimer formation is another dADAR regulatory mechanism (Gallo et al., 2003).

1.3.7 *Drosophila* ADAR null mutants

Drosophila ADAR null mutant flies have been generated using two different approaches. Having identified the dADAR enzyme Palladino *et al.* generated a deletion mutant by site-selected P element mutagenesis followed by excisions of the P element (Palladino et al., 2000b). 13 deletions of *dAdar* were recovered with one allele, 5G1, completely deleting the *dAdar* locus

including all coding sequences. All mutants were conditionally lethal but under ideal growth conditions developed into morphologically normal adults displaying significant behavioural deficits. A translocation of a Y chromosome containing the complete *dAdar* locus to the tip of the X chromosome rescued all phenotypes indicating that the mutant phenotypes observed are the result of null alleles of *dAdar* (Palladino et al., 2000b). The 5G1 allele produced no detectable *dAdar* transcripts and crude extract from *dAdar* null mutant adults showed no A-to-I conversion activity on synthetic dsRNA (Palladino et al., 2000b).

Using a forward genetic approach, Ma *et al.* performed a genetic screen looking for loss-of-function mutations involved in sensitivity to anoxia (Ma et al., 2001). Two mutants were identified, *hypnos-2*^{L} and *hypnos-2*^{P}, that mapped to the same locus with *hypnos-2*^{P} being more sensitive to oxygen deprivation. Analysis of the *hypnos-2*^{P} locus identified the deletion to be within the *dAdar* gene and to abolish the second dsRBM and part of the deaminase domain located in exons 5 and 6. *hypnos-2*^{P} proteins were detected in the *Drosophila* embryo and adults but were smaller than those detected in Canton-S wild-type flies indicating the deletion to be in-frame (Ma et al., 2001). A transgene encoding wild-type dADAR rescued the *hypnos-2*^{P} phenotype indicating that deletion of the second dsRBM and part of the deaminase domain is sufficient to abolish dADAR editing activity (Ma et al., 2001).

dAdar null mutant adult flies present severe behavioural deficits involving slow uncoordinated locomotion, occasional tremors, abnormal body posture, and excessive grooming (Ma et al., 2001; Palladino et al., 2000b). Adult male *dAdar* null flies display extreme mating difficulties and do not initiate any of the normal courtship rituals (Palladino et al., 2000b). A number of editing targets identified in *Drosophila* including *cacophony* (Smith et al., 1996), *paralytic* (Hanrahan et al., 1998), *nap* and *slowpoke* (Hoopengardner et al., 2003) are involved in the courtship song display (Peixoto and Hall, 1998).

The absence of editing activity at these substrates may explain the mating defect seen in male *dAdar* null mutants.

Under optimal conditions *dAdar* null mutants have a life span comparable to wild-type flies but if incubated at 25 °C or in a high-density population they have a higher mortality rate than wild-type flies (Ma et al., 2001; Palladino et al., 2000b). The recovery time from anaesthesia or anoxia is longer and they appear to be extremely sensitive to heat shock, which exacerbates the already severe locomotion and behavioural defects (Ma et al., 2001; Palladino et al., 2000b).

All *dAdar* null mutants display gross anatomical abnormalities within the brain. From a young age the retina is disorganised with photoreceptors extending further to reach the laminar layers (Palladino et al., 2000b). Lesions appear in the central brain, optic lobes and retina in adult flies increasing in number and size with age (Palladino et al., 2000b). Cortical neurons within the medulla, lobula complex and lamina also degenerate (Ma et al., 2001).

Editing activity on the *para* Na^+ channel, *Dmca1A* (*cacophony*) Ca^{2+} channel, *DrosGluCl-α* glutamate-gated chloride channel (Ma et al., 2001; Palladino et al., 2000b), *Dα6* nicotinic acetylcholine receptor subunit (Grauso et al., 2002), and the *Shab* K^+ channel (Bhalla et al., 2004) transcripts is abolished in *dAdar* null mutants indicating that dADAR is the enzyme responsible for site-specific A-to-I editing in all these substrates. Lack of editing of the *para* Na^+ channel has been shown to significantly lower membrane potential in cultured neurons demonstrating a functional electrophysiological consequence of ADAR activity (Ma et al., 2001).

1.3.8 *C. elegans* ADR

Using a differential display screen, Morse *et al.* identified 5 new ADAR substrates in *C. elegans*, 4 of the transcripts are edited in untranslated regions

(UTRs) whilst the fifth is predicted to be a non-coding RNA (Morse and Bass, 1999). The proteins responsible for this activity were identified from a *C. elegans* cDNA library. The two open reading frames were predicted to encode proteins with C-terminal domains similar to the deaminase domain of the ADAR proteins (Hough et al., 1999; Morse and Bass, 1999). *adr-2* was found to encode a protein with one dsRNA binding motif and a highly conserved deaminase domain containing the CDA consensus of $HAE(x)_{41\text{-}58}$ $PCG(x)_{44\text{-}154}$ SCSDK. *adr-1* was shown to have two dsRNA binding motifs and further analysis revealed 5 possible spliceoforms. The deaminase domain of *adr-1* is not as conserved as *adr-2* with a consensus sequence of $DAI(x)_{48}$ $PPC(x)_{42}$ CTADK. This may account for its apparent lower levels of editing activity (Tonkin et al., 2002).

adr-1 was found to be highly expressed in the nervous system throughout worm development and remains present in adult worms. Interestingly, *adr-1* expression was also seen in the developing vulva during morphogenesis of this organ. *adr-2* has been found to be expressed ubiquitously in early embryos (Tonkin et al., 2002).

1.3.9 *C. elegans adr-1* and *adr-2* null mutants

The 10 known targets in *C. elegans* for the ADR proteins are all edited in non-coding regions therefore the physiological function of RNA editing in worms is unclear. Homozygous deletions of *adr-1* and *adr-2* and a double mutant were isolated and editing and behavioural activities were analysed (Tonkin et al., 2002).

All of the mutant worms appeared normal indicating that *adr-1* and *adr*-2 are not essential genes in worms unlike in mammals. *adr-1* mutant worms showed a significant decrease in ADAR activity both *in vitro* and *in vivo* however, *adr-2* and the *adr*-1;*adr*-2 double mutant exhibited no detectable ADAR activity. All of the mutant worms demonstrated defects in

chemotaxis, a measure of the worms olfactory system, with the *adr-1* worms having a relatively mild phenotype when compared to the *adr-2* and the *adr-1;adr-2* mutants. These results suggest that the ADR-2 protein is active on its own but the ADR-1 protein requires ADR-2 for its activity, perhaps the 2 proteins function as a heterodimer.

In addition to the phenotypes described above, *adr-1* mutants and double mutants had defects in their somatic gonad resembling the protruding-vulva (Pvl) phenotype. This corresponds with the expression pattern of *adr-1* seen during morphogenesis of the vulva (Tonkin et al., 2002).

Interestingly, even though the ADRs do not antagonise the RNAi pathway in *C. elegans* when dsRNA is injected, transgenes expressed in somatic tissues of wild-type animals are silenced in the *adr-1;adr-2* mutants with the double mutant having a stronger effect than either of the single *adr* mutants. U to C changes were found in the antisense strand of the transgene in wild-type worms indicating that when present, the ADR proteins deaminate these dsRNAs resulting in I-U mismatches and preventing the dsRNA molecules from initiating the RNAi pathway (Knight and Bass, 2002).

Worms with *rde-1* or *rde-4* mutations resulting in a defective RNAi pathway, rescued the chemotaxis phenotype seen in *adr* mutants (Tonkin and Bass, 2003). This data along with the RNAi phenotype seen in *adr* null worms suggests that the ADAR proteins play a role in monitoring which dsRNAs enter the RNAi pathway.

1.3.10 ADATs

In higher eukaryotes, eight tRNAs (Ala, Leu, Ile, Val, Ser, Pro, Thr, Arg) have been found to carry inosine at the wobble position (position 34) of the anticodon. Eukaryotic $tRNA^{Ala}$ also has inosine at position 37. This is the result of a zinc-mediated hydrolytic deamination reaction where a water molecule is activated by zinc and attacks the carbon-6 of the adenine base to

produce a keto group and the release of ammonia. In the case of $tRNA^{Ala}$, inosine 37 is then methylated by a S-adenosylmethionine (SAM) to form the N^1-methylinosine-37. In absence of SAM inosine 37 containing tRNA accumulates indicating that the deamination reaction is first to occur (Grosjean et al., 1996).

Tad1p (scADAT1) was identified in the yeast *Saccharomyces cerevisiae* as the enzyme that specifically deaminates A37 of $tRNA^{Ala}$ to inosine. It was identified because of its homology to the C-terminus of the ADARs and is not essential for cell viability. Unlike the ADARs, these ADAT enzymes have no dsRBMs but bind tRNA through an as yet uncharacterised RNA binding domain. This binding appears to be dependent on the correct folding of the L-shaped $tRNA^{Ala}$ substrate as well as the local conformation of the anticodon loop. Mutations affecting the three-dimensional structure of the tRNA molecule eliminate the deamination of A37. tRNA editing has also been found to be strictly magnesium dependent (Auxilien et al., 1996). In the absence of Mg^{2+} ions no A-to-I conversion can be seen on tRNA. This is believed to be due to the involvement of Mg^{2+} ions in maintaining the highly conserved tRNA architecture that is recognised by the ADAT enzyme and required for activity. Because of this highly specific substrate requirement it is unlikely that an adenosine in long dsRNAs could be a substrate for Tad1p (Gerber et al., 1998).

The human and *Drosophila* version of this enzyme, hADAT1 and dADAT1 have also been cloned (Keegan et al., 2000; Maas et al., 1999). These proteins have been shown to specifically edit adenosine 37 of $tRNA^{Ala}$ to inosine and are more similar to the ADARs than Tadp1 supporting the notion of a common evolutionary origin for ADARs and ADATs.

In a search of the *Saccharomyces cerevisiae* genome for open reading frames encoding putative deaminases Gerber *et al.* identified the *TAD2* gene. Yeast with one copy of this gene deleted was not viable showing that *TAD2* is an

essential gene. The Tad2p enzyme was found to be involved in the deamination of A_{34} of yeast tRNA and was also found to modify the *E. coli* $tRNA^{Arg}$. This deaminase was found to be a heterodimer consisting of 2 subunits, Tad2p and Tad3p with Tad2p being the catalytic subunit. Like *TAD2*, the *TAD3* gene was also found to be essential indicating that the inosine at the wobble position of tRNAs is a crucial modification in yeast. Tad2p and Tad3p share the deaminase motifs of the cytidine deaminase family suggesting that Tad2p/Tad3p are ancestors of the deaminase domain found in Tad1p and the ADAR family (Gerber and Keller, 1999).

A prokaryotic tRNA that is modified at position 34 has also been identified. *E. coli* $tRNA^{Arg2}$ is modified at the wobble position by the tadA homodimer, converting the A_{34} to inosine. The *tadA* gene has been shown to be essential indicating that modification at the wobble position is essential in prokaryotes also (Wolf et al., 2002).

1.4 Transcripts edited by ADARs

1.4.1 Structural characteristics of pre-mRNA that undergo editing

What elements within a specific transcript dictate that a particular adenosine is edited? Deletion analysis of known pre-mRNA has identified *cis* elements involved in the regulation and determination of what adenosines are edited in a given mRNA sequence.

Studies on the *GluR-B* Q/R editing site demonstrates that editing occurs prior to splicing and that intronic sequences are required (Egebjerg et al., 1994; Higuchi et al., 1993). Deletion analysis identified regions in the downstream intron necessary for RNA editing (Brusa et al., 1995; Egebjerg et al., 1994; Higuchi et al., 1993) (Figure 1.5). None of these regions are present in the intronic sequences of the unedited *GluR-A*, *GluR-C* or *GluR-D* genes

(Egebjerg et al., 1994). Sequence analysis revealed one region, ~310 nucleotides downstream, consisting of 10 nucleotides complementary to the unedited exonic sequence surrounding the Q/R site (Higuchi et al., 1993). Base pairing between this exonic complementary sequence or ECS and the nucleotides around the Q/R site is essential for efficient editing. Substituting 4 of the intron nucleotides to extend the intron-exon duplex and certain paired complementary exonic and intronic substitutions reduced the extent of editing at Q/R site suggesting that a duplex region is necessary but certain nucleotide sequences and structural characteristics are required to ensure accurate and efficient editing of the mRNA molecule (Higuchi et al., 1993).

Formation of this dsRNA duplex structure required for editing is aided by the presence of an imperfect inverted repeat sequence in the proximal portion of the downstream intron (Egebjerg et al., 1994; Higuchi et al., 1993). This inverted repeat is predicted to fold into two hairpins consisting of 3 helical regions and disruption of this structure can decrease or even eliminate Q/R editing (Egebjerg et al., 1994) (Figure 1.5).

Analysis of the Q/R editing sites in GluR5 and GluR6 identified an ECS 1900 nucleotides distal to the editing site (Herb et al., 1996) and deletion of the ECS region of *GluR6* in mice resulted in no Q/R site editing (Vissel et al., 2001) further demonstrating the physiological importance of these non-coding, and sometimes distant regions.

Sequence analysis of the *GluR-B*, *GluR-C* and *GluR-D* R/G sites in exon 13 and the downstream ECS-containing intron 13 demonstrated that the bases in paired helical regions of this dsRNA duplex were conserved whereas nucleotides in non-helical regions varied (Aruscavage and Bass, 2000). Interestingly, this pattern of conservation was not observed in the *GluR-A* gene (Aruscavage and Bass, 2000). This data would suggest an evolutionary pressure by an ADAR-induced mechanism to retain certain intronic sequences thus preserving the structural characteristics of this dsRNA

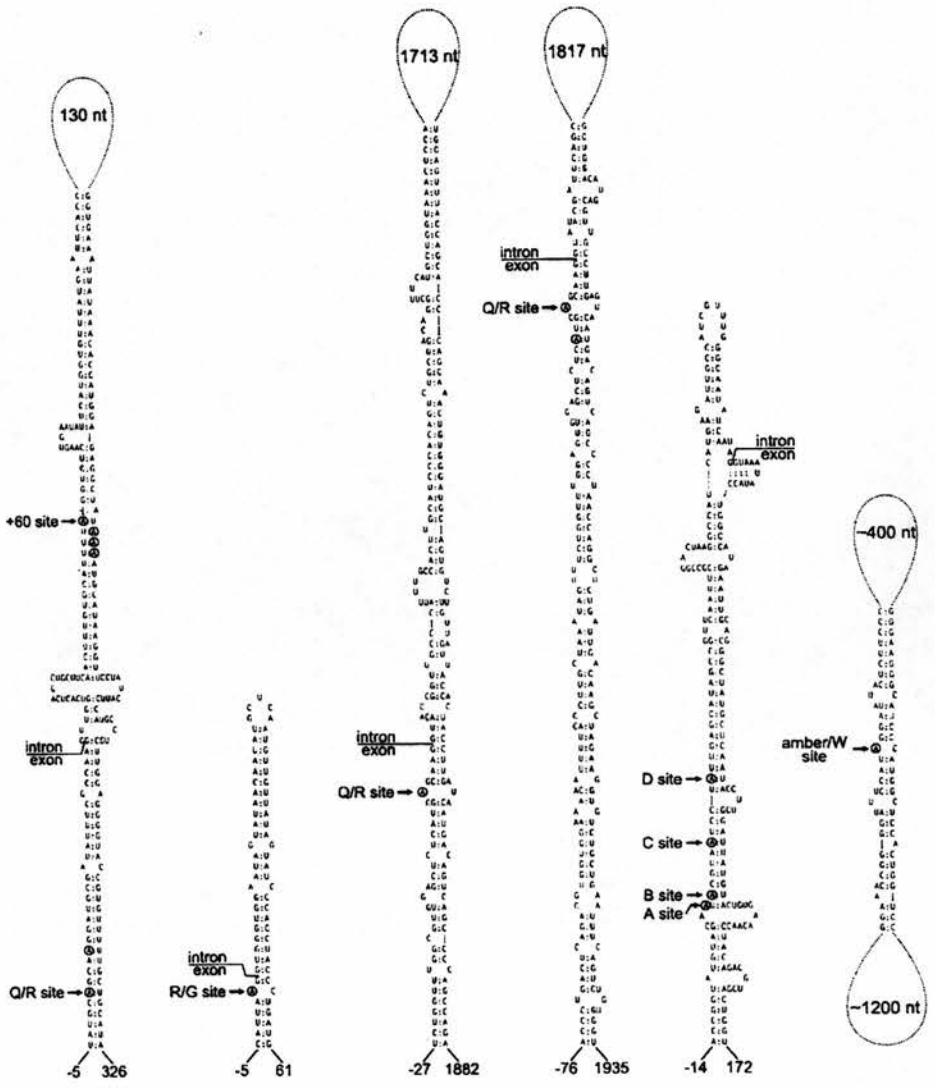


Figure 1.5 Predicted secondary structure of the pre-mRNA transcripts encoding the (from left to right) GluR-B Q/R site, GluR-B R/G site, GluR5 Q/R site, GluR6 Q/R site, 5-HT$_{2C}$R, and HDV amber/W site.
From Chapter 19, Grosjean and Benne, 1998

duplex that facilitate editing at the R/G site.

Not all ECS have that have been identified reside in the downstream intron. Two *D. melanogaster* substrates, the ADAR self-edit site (Palladino et al., 2000a) and the FSP site of the *para* Na^+ channel (Hanrahan et al., 2000) have been shown to have an ECS within the same exon that contains the edited adenosine. Many targets in *D. melanogaster* are conserved in other *Drosophila* species however, the FSP site is not edited in *D. virilis* but sequence analysis would suggest dADAR is edited in *D. pseudoobscura*. Indeed, structural analysis of the *syt1* pre-mRNA substrate in *Drosophila* indicates a specific pseudo knot structure conserved through evolution is responsible for driving editing (Reenan, 2005).

In a perfectly paired dsRNA duplex, ADAR edits multiple adenosines yet in physiological substrates this activity is highly specific, targeting a limited number of particular adenosines. This selectivity is determined by the structural characteristics of the dsRNA substrate. Internal loops act to divide a substrate into smaller duplexed regions that appear as short dsRNA substrates and are deaminated more selectively therefore reducing the number of editing events across the whole molecule (Lehmann and Bass, 1999).

A number of studies have suggested that the efficiency of ADAR editing at a particular adenosine is influenced by its base pairing/mismatch status. Biochemical analysis indicates that when there is an A:C mismatch at the editing site rather than an A:U base pair or A:G/A:A mismatch, ADARs are enzymatically more efficient (Stephens et al., 2000; Wong et al., 2001). However, in naturally occurring substrates this can vary. The adenosine at the Q/R site of GluR-B is base paired yet at the R/G site the adenosine is mismatched (Lomeli et al., 1994) and the Q/R target sites of GluR5 and GluR6 are within a bulge due to a 4 bp mismatch between the exon and ECS

sequences (Herb et al., 1996). It is unclear whether this base pair bias has any significant influence on the efficiency of the ADARs at these different sites.

1.4.2 Glutamate-gated ion channels

The first examples of A-to-I editing in mammals were observed in mRNA transcripts encoding glutamate-gated ion channels in the central nervous system (Sommer et al., 1991). These receptors are involved in fast synaptic neurotransmission and calcium entry through these channels plays an important role in development and in forms of synaptic plasticity that may underlie higher order processes such as learning and memory (Dingledine et al., 1999).

There are three classes of glutamate receptor channel (GluRs); the NMDA (N-methyl–D-aspartate) receptors, the AMPA (α-amino–3–hydroxyl–5–methyl–4-isoxazolepropionic acid) receptors and the kainate receptors (Dingledine et al., 1999). Only non-NMDA receptor transcripts are substrates for RNA editing.

The GluRs consist of four different, sequence-related subunits and the properties of the channel are dependent on which subunits are assembled. Each subunit has four transmembrane hydrophobic regions, TMI- TMIV with TMII forming a re-entrant loop resulting in an extracellular N-terminus and intracellular C-terminus. The large loop between TMIII and TMIV is exposed to the cell surface and forms part of the ligand binding domain (Dingledine et al., 1999).

1.4.2.1 AMPA receptors

The AMPA receptors are assembled from a combination of the GluR-A, GluR-B, GluR-C and GluR-D subunits into homo- or heteromeric channels. Studies have shown that the GluR-B subunit is dominant in determining the functional properties of a heteromeric AMPA receptor channel (Jonas and

Sakmann, 1992). Native AMPA receptor channels are impermeable to calcium and this impermeability is determined by the GluR-B subunit (Burnashev et al., 1992). The GluR-B mRNA is edited at a site in the channel-forming TMII region resulting in a single amino acid change from glutamine (Q) to arginine (R) (Sommer et al., 1991). The positive charge of arginine at this Q/R site contributes to low calcium permeability suggesting that this residue is critical in determining AMPA receptor properties (Burnashev et al., 1992) (Figure 1.6).

This glutamine to arginine amino acid change is the result of an enzymatic conversion of the adenosine base in the glutamine CAG codon to inosine (Rueter et al., 1995). The inosine is translated as though it were guanosine, resulting in an arginine residue (CGG). This deamination was found to be mediated by a dsRNA adenosine deaminase activity (Hurst et al., 1995; Melcher et al., 1995; O'Connell et al., 1997) and the presence of a downstream intronic sequence, the ECS, was shown to be essential (section 1.5.2) (Brusa et al., 1995; Egebjerg et al., 1994; Higuchi et al., 1993).

ADAR2 was identified as the enzyme responsible for Q/R editing (Higuchi et al., 2000; Melcher et al., 1995). ADAR2 null mice were prone to seizures and died young however, the insertion of the edited allele of GluR-B rescued this phenotype (section 1.4.2) (Higuchi et al., 2000). Overexpression of the genomically encoded GluR-B (Q) subunit results in an increase in AMPA receptor conductance and epileptic seizures and premature death (Brusa et al., 1995; Feldmeyer et al., 1999) further highlighting the physiological importance of Q/R editing. Interestingly, it has been shown that the GluR-B (Q) subunit is not essential for brain development and function (Kask et al., 1998) though a compensatory mechanism is believed to be involved when GluR-B is missing.

Greger *et al.* demonstrated that the GluR-B subunit resides in an intracellular pool within the endoplasmic reticulum (ER) and that this ER retention is

controlled by the arginine residue at the Q/R editing site (Greger et al., 2002). GluR-B (R) subunits are largely dimeric and are retained in the ER unlike unedited GluR-B (Q) subunits which readily tetramerize and are rapidly exported from the ER and trafficked to the cell surface (Greger et al., 2003; Greger et al., 2002). This suggests that the arginine at the Q/R site is a unique ER retention signal and alters a region of the subunit involved in receptor assembly thus restricting the incorporation of GluR-B (R) subunits in AMPA receptor tetramers (Greger et al., 2003; Greger et al., 2002).

These results suggest that in studies where Q/R editing is abolished the phenotypes seen may not only be due to an increase in calcium permeability of each individual channel but also changes in AMPA receptor stoichiometry and number.

Lomeli *et al.* identified another RNA editing site in exon 13 of GluR-B resulting in an arginine to glycine codon switch (Lomeli et al., 1994). Termed the R/G site, this induced amino acid change was found to also occur in the GluR3 and GluR4 transcripts (Lomeli et al., 1994).

The R/G site is located immediately upstream of the alternatively spliced flip-flop cassette and the extent of editing at this site varies in a flip-flop splice-dependent manner (Lomeli et al., 1994). A conserved region critical for R/G editing was identified in the proximal portion of intron 13 of the *GluR-B*, *GluR-C* and *GluR-D* genes and this region has been deemed the ECS (Lomeli et al., 1994). R/G site editing of the GluR subunits results in slower AMPA receptor desensitization and an increase rate of recovery (Lomeli et al., 1994). Rapidly recovering receptors are able to respond to repeated glutamate stimulation thus allowing the AMPA receptors to transmit signals from fast stimuli.

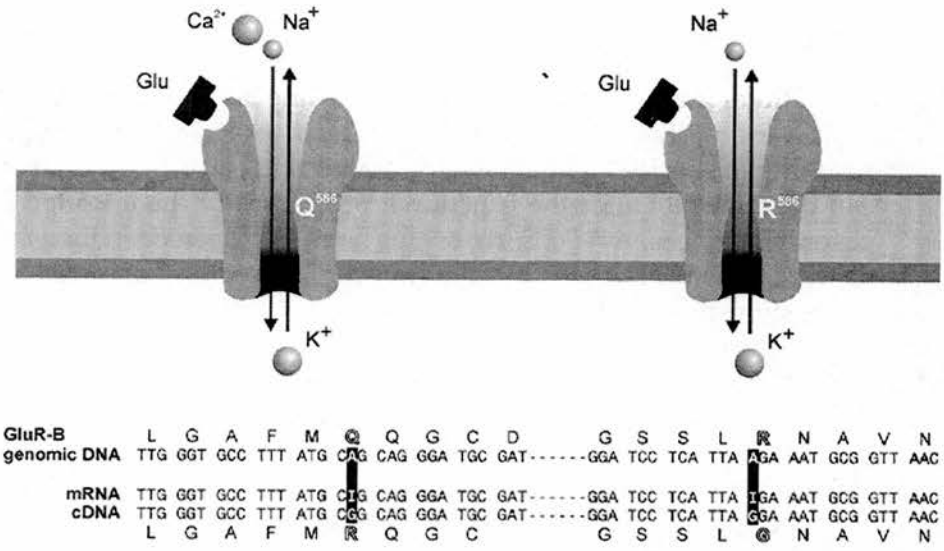


Figure 1.6 Schematic diagram of the heteromeric AMPA receptors indicating ion-permeation properties of channels containing non-edited (Q586) or edited (R 586) GluR-B receptor subunits. Nucleotide and predicted amino acid sequence alignments between GluR-B genomic, mRNA and cDNA sequences; nucleotide discrepancies and predicted alterations in amino acid sequence are indicated in inverse and open lettering, respectively. (Taken from RNA Editing, 2000 Ed. B.L. Bass)

1.4.2.2 Kainate receptors

Kainate receptors can be generated from the heteromeric assembly of the subunits GluR5, GluR6, KA-1 and KA-2 (Dingledine et al., 1999). Like the AMPA receptor GluR-B subunit, GluR5 and GluR6 transcripts are edited in a critical position (Q/R site) of the putative channel-forming segment TMII (Egebjerg and Heinemann, 1993; Köhler et al., 1993). However, in contrast with AMPA receptor channels, the presence of glutamine at the Q/R site results in channels with low calcium permeability, whereas an arginine determines higher calcium permeability (Egebjerg and Heinemann, 1993; Köhler et al., 1993).

Unlike the AMPA receptors, calcium permeability is not solely controlled by Q/R editing. Sequence analysis of GluR6 identified two other RNA editing sites within the TMI region that change an isoleucine residue to valine (I/V site) and a tyrosine residue to a cysteine (Y/C site). If the TMI sites are unedited, calcium permeability is less influenced by the presence of either glutamine or arginine at the Q/R site of TMII thus Q/R editing in kainate receptors is not the only determinant of calcium permeability (Köhler et al., 1993).

The GluR5 and GluR6 editing sites are not edited to 100% like the GluR-B Q/R site but instead different edited variants are co-expressed in a single cell (Ruano et al., 1995). Kainate Q/R site editing is mediated by a dsRNA adenosine deaminase activity (Herb et al., 1996) and is developmentally regulated (Lee et al., 2001). An ECS has been identified in GluR5 and GluR6 located 1900 nucleotides from the Q/R site (Herb et al., 1996) and is essential for kainate Q/R editing (Vissel et al., 2001). Mice with an ECS deletion in the *GluR6* gene express only the GluR6 (Q) subunit (Vissel et al., 2001). Phenotypically they display no behavioural deficits but have increased

seizure vulnerability (Vissel et al., 2001). This suggests that RNA editing at the Q/R site of the kainate GluR5 and GluR6 subunits may play a role in seizure susceptibility and perhaps be potential targets for antiepileptic drugs.

1.4.3 Serotonin (5-HT_{2C}) receptor

The 5-HT_{2C} receptor is a seven-transmembrane-spanning G-protein coupled receptor that undergoes RNA editing at 5 distinct sites located in the coding region of the second intracellular loop, a region of the protein involved in the coupling of receptors to G protein-mediated signalling cascades (Burns et al., 1997; Fitzgerald et al., 1999) (Figure 1.7). Editing patterns at sites A-E occur in a tissue-specific manner and can result in, amongst others, a VGV or VSV isoform of the 5-HT_{2C} receptor, both of which exhibit a reduced coupling ability to the G protein (Burns et al., 1997).

The serotonin receptor has been shown to be constitutively active, that is, it can couple to the G protein α subunit in the absence of an agonist. The basal phosphoinositide hydrolysis activity of the rat non-edited 5-HT_{2C} INI isoform was substantially higher than that of the VGV isoform yet when treated with serotonin both isoforms have the same magnitude of increase in inositol phosphate formation. These data indicate that by editing the 5-HT_{2C} transcript the constitutive activity of the receptors is silenced not its ability to couple to G proteins (Niswender et al., 1999). Further studies on the human receptor indicate that changes in agonist high affinity binding are also seen in edited isoforms and contribute to the overall decrease in downstream effects (Fitzgerald et al., 1999). This reduction in agonist potency depends on the intrinsic activity of the ligand with full agonists being most affected by editing and antagonists showing little effect.

RNA editing occurs specifically in the region of the 5-HT_{2C} receptor that interacts with the G protein α subunit. Price *et al.* have demonstrated that the unedited INI isoform is capable of interacting with the α subunit of both the

G_q and G_{13} proteins whereas the edited VGV and VSV isoforms are unable to activate the G_{13} protein (Price et al., 2001). This results in a potential change in 5-HT_{2C} receptor downstream effects as the G_{13} protein does not interact with the adenyl cyclase or phospholipase pathways but instead has been found to alter the actin cytoskeleton arrangement via Pho-dependent mechanisms (Gohla et al., 1999). Further studies have shown that unlike the edited VGV isoform, the unedited INI isoform is capable of activating Rho-GTPases via G_{13} and the phospholipase D pathway, a pathway implicated in downstream responses such as actin cytoskeletal rearrangement and exocytosis (McGrew et al., 2004).

A number of contradictory pathological studies have implicated changes in the RNA editing patterns of the 5-HT_{2C} receptor in a number of psychiatric disorders including schizophrenia (Dracheva et al., 2003; Iwamoto and Kato, 2003; Sodhi et al., 2001), depression (Gurevich et al., 2002b; Iwamoto and Kato, 2003; Iwamoto et al., 2005; Yang et al., 2004) and suicide (Gurevich et al., 2002b; Iwamoto and Kato, 2003; Niswender et al., 2001; Schmauss, 2003).

Using post-mortem samples from schizophrenic patients Sodhi *et al.* found an overall reduction in RNA editing of the 5-HT_{2C} receptor with higher levels of the basally more active unedited INI isoform (Sodhi et al., 2001). This however, was contradicted by 3 other groups who found no such changes (Dracheva et al., 2003; Iwamoto and Kato, 2003; Niswender et al., 2001). These different outcomes may partly be explained by the sample sizes of each study with Sodhi *et al.* using a sample size thought to be too small to produce any statistically significant results in such a complex editing substrate. In addition, the mix of male and female subjects and the presence or absence of drug treatment may also play a role in the differing results seen between all the studies. Due to the complexity of schizophrenia and the wide range and severity of symptoms seen by clinicians it is difficult to ascertain the exact role RNA editing of the serotonin 5-HT_{2C} receptor plays in the pathology of schizophrenia.

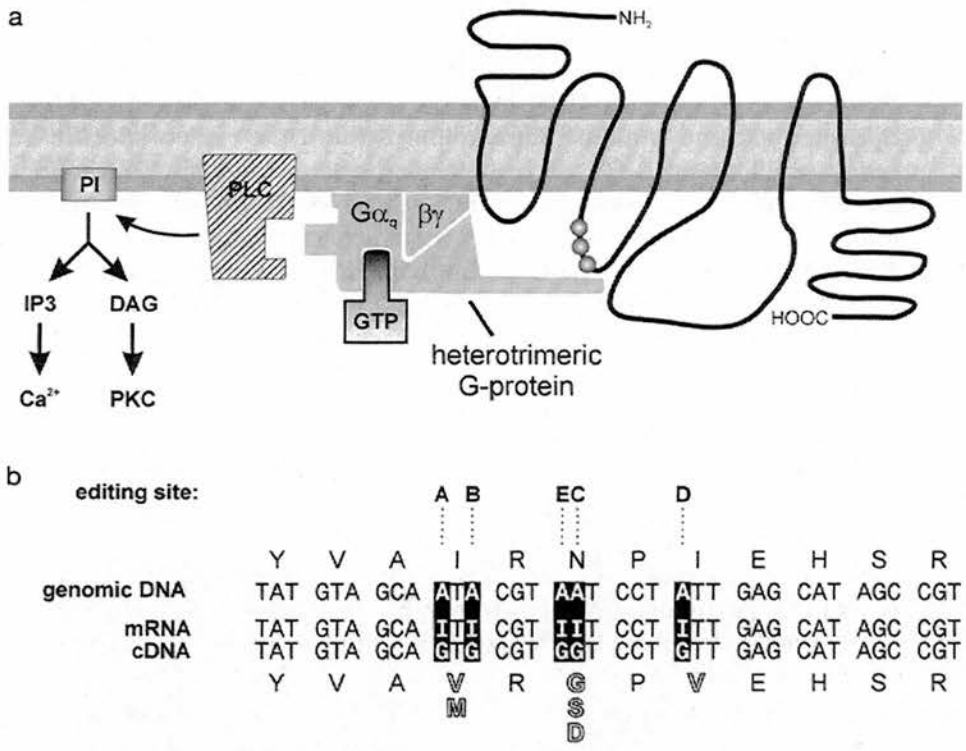


Figure 1.7 RNA Editing of 5-HT2CR transcripts.

A. A schematic representation of the predicted topology for the 5-HT2C receptor is presented.

B. Nucleotide and predicted amino acid sequence alignments between 5-HT2CR genomic, mRNA and cDNA sequences. (Taken from RNA Editing, 2000 Ed. B.L. Bass)

One factor highlighted by the Niswender *et al.* study was the elevated levels of editing at the A-site seen in patients who had committed suicide regardless of their underlying psychiatric disorder (Niswender et al., 2001). This result was replicated in a later study (Iwamoto and Kato, 2003) suggesting a link between the change in serotonin-mediated signalling due to altered editing efficiencies at the A-site and the particular mental state associated with suicide.

In contrast with the studies mentioned previously, Gurevich *et al.* found an increase in editing at the E-site, a decrease at the D-site with the C-site showing a trend towards an increase in suicide victims with a history of major depression. The treatment of mice with Prozac (fluoxetine), a serotonin-selective reuptake inhibitor, resulted in the complete opposite changes in editing patterns than those seen in suicide victims, suggesting a serotonin-dependent site-specific regulation of 5-HT_{2C} RNA editing (Gurevich et al., 2002b). In fact, further studies indicate that serotonin-depleted animals show a decrease in editing at sites C and E resulting in increased expression of more serotonin-sensitive 5-HT_{2C} receptor isoforms (Gurevich et al., 2002a). When treated with a partial agonist an increase in RNA editing at the E-site is observed and hence the expression of 5-HT_{2C} isoforms less responsive to serotonin activation. This suggests a role for RNA editing in maintaining 5-HT_{2C} receptor activation within an optimal range in response to changes in serotonin levels (Gurevich et al., 2002a).

These complex and conflicting data indicate that RNA editing may play a role in regulating fundamental properties of neuronal cells including polarity, migration and transformation via its A-to-I editing activity on the serotonin 5-HT_{2C} receptor with these changes potentially resulting directly or as seems more likely, indirectly, in the pathology and prognosis of a number of major psychiatric disorders.

1.4.4 Transcripts edited in *Drosophila*

1.4.4.1 The *4f-rnp* gene

The first transcript to be edited in *Drosophila* was encoded by the *4f-rnp* gene. Comparisons of genomic DNA and cDNA revealed multiple A-to-G conversions in transcripts isolated from *Drosophila* adult head (Petschek et al., 1996). The biological function of the *4f-rnp* gene is unknown but transcripts are particularly abundant in the central nervous system and deletions of the gene result in lethality during the larval period indicating an essential role for *4f-rnp* (Petschek et al., 1996). Sequence analysis indicate it is a novel gene with the non-edited transcript encoding a 943 amino acid protein containing a C-terminal RNA binding domain (Petschek et al., 1997). Editing events at some of the 263 sites identified result in altered splicing patterns and the translation of a number of isoforms (Petschek et al., 1997).

Peters *et al.* identified the *sas-10* gene as the source of the antisense RNA transcript that complements *4f-rnp* forming the dsRNA substrate required for A-to-G modification (Peters et al., 2003). The *4f-rnp* and *sas-10* genes are adjacent, located on opposite DNA strands and convergent transcription results in long readthrough transcripts of *sas-10* complementary to *4f-rnp* transcripts. The expression of this *sas-10* transcript is developmentally regulated appearing at the late embryo stage of *Drosophila* development and accompanied by an increase in editing and decline in *4f-rnp* transcript levels. This suggests degradation or nuclear retention of the *4f-rnp* transcripts but like other transcripts regulated by a similar antisense/RNA editing mechanism (Section 1.5.8.3) the role this system plays in gene regulation is yet to be fully explained.

1.4.4.2 The Calcium channel α1 subunit, *Dmca1A*

The α1 subunit of the *Drosophila* voltage-sensitive calcium channel (*Dmca1A*) was found to map to the locus identified in the *cacophony* (*cac*) courtship and

night-blind-A visual (*nbA*) mutations on the X chromosome (Smith et al., 1996). The Dmca1A protein consists of four homologous repeats each composed of six putative α-helical transmembrane domains and forms the calcium conductance pore of the voltage-gated calcium channel. The fourth transmembrane domain of each repeat has positively charged residues every 3 or 4 amino acids and is thought to be the voltage sensor whilst the link between the fifth and sixth transmembrane regions forms part of the pore (Peixoto et al., 1997).

11 putative editing sites were identified in *Dmca1A* transcripts resulting in 10 amino acid changes (Peixoto et al., 1997; Smith et al., 1996). Putative ECS's have been identified for six of the *Dmca1A* editing sites and all contain a 9-nucleotide region of complementarity perfectly centred on the edited adenosine (Smith et al., 1998). Palladino *et al.* generated a *dADAR* null mutant in which no A-to-G changes were detected in the *Dmca1A* transcripts demonstrating that the dADAR protein is responsible.

Along with the various alternatively spliced regions, editing at these 11 sites provides another level of variability to the properties of the translated protein and resulting in potentially 125,000 different isoforms.

1.4.4.3 The *para* voltage-gated Na^+ channel

The *Drosophila para* locus was found to encode a voltage-sensitive sodium channel in *Drosophila* neurons similar in structure to the α subunit of vertebrate sodium channels (Loughney et al., 1989). Like the *Dmca1A* channel, the *para* channel consists of four hydrophobic homology domains. Each of these repeats comprises six transmembrane α-helical structures with the fourth segment having a positively charged residue every three amino acids and representing the voltage sensor.

Analysis of *para* cDNA initially identified six editing sites, five of which result in an amino acid change (Hanrahan et al., 1998). An A-to-G change at

the FSP and SSP editing sites eradicate *Fsp*I and *Ssp*I restriction enzyme sites and result in Q/R and K/R amino acid changes, respectively whilst editing at the SFC site generates a *Sfc*I restriction enzyme site and a N/S amino acid change. These sites undergo editing independent of each other and alternative splicing with levels varying throughout development and an obvious increase apparent between embryo and adult (Hanrahan et al., 1998). Complementary regions were identified in downstream introns for the SSP and SFC sites with editing at both appearing to be evolutionarily maintained (Hanrahan et al., 1998; Hanrahan et al., 2000). Unsurprisingly the exonic sequences of the *para* gene are conserved between both *D. melanogaster* and *D. virilis* however, the intronic sequences incorporating the ECS for the SSP and SFC editing sites were also highly conserved (Hanrahan et al., 2000).

Editing at the FSP site was not evident in *D. virilis* (Hanrahan et al., 1998; Hanrahan et al., 2000) and a search for the ECS indicated that it may lie within exonic sequences upstream of the editing site. A transgene containing the intron upstream of the edited exon and part of the upstream exon was efficiently edited when injected into flies indicating that the structure necessary for A-to-I editing is constructed from these sequences (Hanrahan et al., 2000).

The 12-nucleotide T/M editing region contains three A-to-G changes introducing *Taq*I and *Mn*II restriction enzyme sites and resulting in a Q/R and N/D amino acid changes (Hanrahan et al., 1998). Sequence analysis indicates that each of these three events occur independently of each other and are evolutionarily conserved between *D. melanogaster* and *D. virilis* (Hanrahan et al., 1998).

Phenotype similarities between the *paralytic-temperature sensitive, para*^{ts}, and the *maleless-no action potential temperature sensitive mle*^{napts} null mutants suggested a link between the wild-type functions of both genes. Further

analysis demonstrated that the lack of RNA helicase function in the *mle*^{napts} null mutants prevent the resolution of the secondary structure around the *para* T/M editing site resulting in appropriate splicing causing exon-skipping of the T/M containing exon (Reenan et al., 2000). The resulting phenotype demonstrates the importance of the editing events at the T/M sites for normal function of the *para* voltage-gated Na+ channel.

Another four editing sites have since been identified in the *para* transcript, three of which incorporate amino acid changes (Palladino et al., 2000b). This brings the total number of editing sites within this transcript to ten and alongside the ten alternative splicing locations; the potential to generate a vast number of isoforms.

1.4.4.4 Other *Drosophila* neuronal targets

Two other classes of ion channel have also been identified as targets of A-to-I RNA editing. The *DrosGluCl-α* subunit of the inhibitory glutamate-gated chloride channel has 5 known editing sites (Semenov and Pak, 1999) and the *Dα6* nicotinic acetylcholine receptor (nAChR) subunit transcript has seven editing sites resulting in amino acid changes in the ligand-binding domain of the receptor (Grauso et al., 2002). Sequence analysis indicated that four of the sites seen in *Drosophila Dα6 nAChR* were also edited in the moth *Heliothis virescens* suggesting an evolutionary conserved function for these editing events (Grauso et al., 2002). A-to-I changes were not detected in *DrosGluCl-α* nor *Dα6* transcripts isolated from *dAdar* null mutant flies indicating they are substrates of the wild-type dADAR enzyme (Palladino et al., 2000b). Comparative sequence analysis of a selection of genes associated with nervous system function in *Drosophila melanogaster* and *pseudoobscura* identified 16 new neuronal targets (Hoopengardner et al., 2003).

Recently Bhalla *et al.* identified a new target, the *Shab* gene a member of the K_v2 gene subfamily of voltage-gated K^+ channel α subunits (Bhalla et al.,

2004). The comparative genetics study between the *melanogaster* and *pseudoobscura* genomes failed to identify *Shab* because it had not been cloned in *pseudoobscura*. Two adjacent adenosines are edited with the 5′ adenosine in the degenerate position of a valine codon and silent and the 3′ adenosine analogous to the editing site in human $K_v1.1$ K^+ channel introducing an I/V site to the *Shab* K^+channel. Editing at this site is attributed to the dADAR enzyme, as this modification is not seen in *dAdar* null mutant flies (Bhalla et al., 2004).

That all the known targets, apart from *4f-rnp* and *dAdar* itself, are involved in the fly nervous system is not surprising when considering the behavioural and locomotion defects seen in the null mutant flies. Interestingly, all the neuronal substrates identified are involved in rapid chemical or electrical neurotransmission suggesting that one of the main roles of the ADAR enzyme in *Drosophila* is the regulation of ion channel function.

1.4.5 *C. elegans* edited transcripts

ADAR activity in *C. elegans* was first identified using a differential display screen looking for A to G changes between genomic DNA and mRNA. Initially 5 new substrates were identified, three mRNAs of unknown function with editing sites in the 3′ or 5′ UTRs, a non-coding RNA, and the *pop-1* transcription factor with the deamination sites in its 3′ UTR (Morse and Bass, 1999). A further 5 substrates were identified in a repeat of the screen with all the editing sites again located in non-coding regions. Interestingly, three of these substrates were identified as mRNAs important for nervous system function (Morse et al., 2002).

Structural analysis of the 10 substrates showed they all had the potential to form stable stem-loop structures with the base-pairing involving inverted repeats found within the mRNA molecules. Multiple A to G changes were found in each substrate with those substrates predicted to be almost

completely double-stranded have many more sites than the substrates containing mismatches, loops and bulges (Morse et al., 2002).

1.4.6 Viral transcripts

1.4.6.1 Hepatitis Delta Virus

Hepatitis Delta Virus (HDV) is a subviral pathogen that requires the hepatitis B virus in order to propagate. HDV particles consist of the HDV genomic RNA, the hepatitis delta antigen (HDAg) and a viral envelope made from the hepatitis B surface antigen. The ~1680 nucleotide closed circular RNA genome has a single open reading frame coding for the HDV antigen protein (Lai, 1995). A-to-I RNA editing by ADAR1 (Jayan and Casey, 2002b; Polson et al., 1996) at position 1012 of the antigenomic RNA, the amber/W site, converts the UAG amber termination codon to a UGG tryptophan codon allowing a long (HDAg-L) and a short form (HDAg-S) of the HDAg protein to be translated (Casey and Gerin, 1995) (Figure 1.5). The HDAg-S protein is essential for RNA replication whilst the HDAg-L protein has been shown to inhibit RNA replication and is essential for viral packaging into the hepatitis B surface antigen envelope (Lai, 1995).

During viral replication, editing at the amber/W site takes place in the nucleus by the constitutively expressed short form of ADAR1 (Wong and Lazinski, 2002). When overexpressed in cultured cells both ADAR1 and ADAR2 have been shown to edit the HDV antigenome (Jayan and Casey, 2002a; Sato et al., 2001; Wong et al., 2001). Increased expression of ADAR proteins results in a strong inhibition of HDV RNA replication, likely due to increased editing at the amber/W site resulting in higher levels of HDAg-L that in turn inhibits viral replication. However, editing at other non-amber/W sites is also seen resulting in other HDAg variants that have been shown to have a dominant negative effect on RNA replication (Jayan and

Casey, 2002a). Therefore levels of ADAR expression play a role in the regulation of HDV antigenome editing.

It is clear that A-to-I editing of the HDV antigenome must be regulated and highly specific for the amber/W site to ensure successful viral replication. If RNA editing is non-specific and occurs at sites on the antigenome other than the amber/W site then these mutations would be carried into the viral RNA genome. Furthermore if editing of the amber/W site is unregulated then only HDAg-L protein would be produced resulting in the inhibition of RNA replication and no propagation of HDV.

Studies have indicated that editing of the HDV antigenome can occur at other positions but around 90% of the time it is only the amber/W site that is edited. This editing can occur independently of viral replication and the HDAg protein however the presence of the HDAg protein has been shown to inhibit editing of the amber/W site in a concentration dependent manner (Polson et al., 1998). The ability of HDAg to bind to the HDV RNA genome and the requirement of a 75 amino acid segment encompassing an RNA binding domain for this inhibitory activity indicates that HDAg may inhibit editing by binding to or near to the amber/W site.

1.4.6.2 A-to-G hypermutation in RNA viruses

A number of RNA viruses have been reported to accumulate A-to-G hypermutations during their life cycle (Cattaneo et al., 1988b; Hajjar and Linial, 1995; Martinez and Melero, 2002; O'Hara et al., 1984). Measles virus (MV), a negative strand RNA virus is one of the best studied. Normally causing an acute infection in children, on rare occasions measles virus can result in a persistent infection leading to subacute sclerosing panencephalitis (SSPE) or measles inclusion body encephalitis (MIBE). This persistence is thought to be due to defective matrix protein (M protein) expression, whose

main function is viral assembly, resulting in lack of viral protein budding (Hall et al., 1979).

The defective expression of MV genes has been correlated to a number of mutations detected throughout the genome in MV genes isolated from SSPE and MIBE cases (Cattaneo et al., 1988a; Cattaneo et al., 1988b; Wong et al., 1989). Many of these mutations result in amino acid changes altering protein sequence and reading frames. Interestingly, a biased hypermutation was identified in the matrix gene where around 50% of the U bases were changed to C (Cattaneo et al., 1988b). These mutations in particular have been identified in a number of measles virus strains resulting in SSPE and are believed to disrupt M protein conformation likely abolishing protein function, altering epitopes and resulting in non-productive infection (Wong et al., 1991). Hypermutation in the viral genome of human respiratory syncytial virus (HRSV) has also been shown to alter protein epitopes and contribute to viral genetic variation (Martinez et al., 1997; Martinez and Melero, 2002).

After the discovery of a dsRNA unwinding activity with the ability to convert adenosine residues to inosine (Bass and Weintraub, 1988) it was postulated that the multiple U-to-C changes seen in the MV matrix gene could be due to A-to-I mutations (Bass et al., 1989). During viral transcription and replication the formation of double stranded regions of RNA involving the plus and minus strands is possible. These regions could be modified by this editing activity. A-to-I changes on the minus strand would result in U-to-C changes in the plus strand and if this plus strand acts as a template for replication, the subsequent negative-stranded genome would contain A-to-G mutations. In cases of persistent infection these mutations could be positively selected in order to inactivate matrix protein function. It should be noted however, that even though ADAR activity has been identified as the enzyme responsible for A-to-G transitions in other RNA viruses (Hajjar and Linial, 1995; Martinez and Melero, 2002) its

involvement in MV hypermutation has yet to be shown (Horikami and Moyer, 1995).

1.4.6.3 Nuclear retention of hyperedited viral transcripts

Throughout the life cycle of polyoma virus, like many other viruses, there is a strict temporal regulation of early- and late-strand transcript levels. Late in polyoma virus infection nuclear antisense RNA from the late strand reduces the levels of early strand transcripts available for translation. Many of these early strands contain adenosine to inosine base modifications and are retained in the nucleus (Kumar and Carmichael, 1997) suggesting that the subsequent dsRNA structures formed when the antisense RNA binds to the early strands are acting as substrates for the dsRNA editing activity by ADAR. This nuclear retention of the modified polyoma early transcripts has since been attributed to a specific inosine-induced retaining complex that anchors hyperedited mRNAs to the nuclear matrix preventing their export to the cytoplasm and subsequent translation (Zhang and Carmichael, 2001).

The bi-directional transcription of the *Xenopus laevis* basic fibroblast growth factor (bFGF) produces complimentary sense and antisense transcripts a small fraction of which undergo multiple A-to-I modifications during maturation of the oocyte (Kimelman and Kirschner, 1989). This hypermutation has been attributed to cellular ADAR activity (Saccomanno and Bass, 1999) and is similar to the extensive A-to-G changes observed in the *Drosophila* gene *4f-rnp* (Section 1.5.4.1) (Peters et al., 2003; Petschek et al., 1996; Petschek et al., 1997). The biological role of this hypermutation was initially thought to trigger an inosine-induced degradation of transcript and studies identified a mammalian ribouclease specific for inosine-containing double-strand RNA (I-RNase) (Scadden and Smith, 1997). This RNase has recently been identified as Tudor-SN, a component of the RISC complex (Scadden, 2005). However, the bFGF transcript was shown to be stable and persist after maturation indicating another possible purpose for this activity.

1.4.7 Other ADAR targets

1.4.7.1 Disease-related Targets

A number of transcripts have been identified that are edited by the ADAR enzymes in the context of particular diseases. It is unlikely that these editing events cause the disease but they may contribute to the overall pathology. One such example has been uncovered in the cells of patients with acute myeloid leukaemia (AML). A novel mRNA species of the PTPN6 gene found to be derived from aberrant splicing within the N-SH2 domain has been discovered (Beghini et al., 2000). PTPN6 is a SH2 domain-containing tyrosine phosphatase mainly expressed in haematopoietic cell development, proliferation and receptor mediated mitogenic signalling pathways. Sequence analysis of PTPN6 transcripts isolated from the cells of AML patients revealed a number of A-to-G conversions mainly occurring at A7866, a putative branch site (Beghini et al., 2000). Editing at this site has been shown to abrogate splicing resulting in the retention of intron 3 leading to nonsense translation and a predicted inactive truncated protein. Levels of the aberrantly spliced mRNA were lower in samples from patients in remission than at the point of diagnosis suggesting the involvement of post-transcriptional PTPN6 processing in leukaemogenesis (Beghini et al., 2000).

A similar aberrant editing event has been identified in the human endothelin B (ETB) receptor gene known to be involved in Hirschsprung disease (HSCR), a congenital intestinal disorder characterised by the absence of ganglionic cells in the distal portion of the intestine (Tanoue et al., 2002). Mutant ETB transcripts lacking a 134 bp nucleotide sequence corresponding to exon 5 were obtained from lymphoblastoid cells from HSCR patients. Some of these splice variants also contained an A-to-G change at position 950 in exon 4 resulting in a glutamine to arginine amino acid change. ADAR1 and ADAR2 were both shown to edit position 950 with ADAR2 being more efficient. Interestingly, this editing event was only seen in the splice variants

lacking exon 5 suggesting that in this instance, the omission of exon 5 suitably alters the structure of the mRNA creating a viable editing site. Western blot analysis also revealed no protein present corresponding to the predicted translated product of this splice variant indicating that the mRNA or protein may be degraded (Tanoue et al., 2002).

Systemic lupus erythematosus (SLE) is an autoimmne disorder characterised by abnormal functions of T-lymphocyte immune effector cells. Changes in signal transduction pathways are thought to contribute to the SLE phenotype. An abnormal cAMP/protein kinase A pathway has been identified in a large number of SLE patients. This defect is due to the deficient activity of the type 1 protein kinase A caused by reduced amounts of both its α and β regulatory subunits (RIα and RIβ). A large number of mutations have been discovered in the RIα transcripts of SLE patients, many of which are T-to-C or A-to-G changes clustered in GAGAG motif and CT repeat "hotspots" (Laxminarayana et al., 2002). ADAR1 expression in SLE T cells was increased 3.5 times when compared to wild-type suggesting that A-to-I RNA editing may play a role in the mutagenesis seen in SLE (Laxminarayana et al., 2002). This overexpression of ADAR1 is thought to be due to the high circulating concentrations of interferon α in SLE patients (Hooks et al., 1979). High levels of ADAR1 protein could explain the aberrant editing of the RIα transcript however, it is unclear whether this editing is responsible for the lower levels of RIα transcript. It is also possible that other transcripts may be edited and so could contribute to the pathology of SLE.

1.4.7.2 A-to-I Editing within Repetitive Elements

By screening poly(A^+) mRNA from human brain for inosine-containing transcripts, 19 new targets for ADAR editing were discovered (Morse et al., 2002). Of the 9 transcripts that were identified from databases, the editing site was located within repetitive elements such as Alu repeats embedded

within non-coding regions of the mRNA including introns and 3′ UTRs (Morse et al., 2002).

More recently many groups have utilised the powerful tool of bioinformatics to identify other ADAR targets within the human transcriptome. Levanon *et al.* aligned human ESTs and cDNAs to genomic sequence lowering possible mismatches by discarding all single-nucleotide repeats longer than 4 bp and removing 150 bps from both ends of the sequence. SNPs of genomic origin were removed and 50-nucleotide windows containing 5 or more mismatches were discarded. Those with 4 or more identical sequential mismatches were retained reflecting the many sequential editing sites that have been previously documented (Morse et al., 2002). The application of these strict criteria to the data resulted in mainly A-to-G mismatches being identified. The same approach looking for G-to-A mismatches yielded only 242 sites indicating there are a low number of false positives attributable to sequence error, SNPs and mutations. Using this method they identified 12,723 putative editing sites belonging to 1,637 different genes, 26 of which were verified by sequences analysis of DNA and RNA samples from up to five tissues (Levanon et al., 2004). Reflecting the results of Morse *et al.* the majority of the sites identified in this screen are located in noncoding regions such as 5′ UTRs, 3′UTRs and introns with 92% of these sites embedded in *Alu* repeat elements (Levanon et al., 2004). This obvious preference for sites within *Alu* repeats may be due to the long paired RNA structures that are likely formed between repetitive elements. However, Kim *et al.* have further documented high levels of A-to-I editing in *Alu* repeats when compared to other repeat elements such as LINE-1, LTR and DNA transposon families (Kim et al., 2004). This suggests a distinct preference for *Alu* repeats and perhaps another potential role for ADARs in modulating the function of these repetitive elements.

1.4.7.3 RNA editing of miRNA precursor molecules

Micro RNAs (miRNAs) are small functional RNA molecules involved in gene regulation (Bartel, 2004). Luciano *et al.* have recently identified a miRNA precursor, pri-miRNA22 that is edited at a number of sites in both humans and mice (Luciano et al., 2004). The ribonuclease III-like enzyme Drosha processes these long pri-miRNA transcripts in the nucleus into ~60-70 nucleotide pre-miRNAs which are then further cleaved by Dicer into mature 20-22 nucleotide miRNAs (Bartel, 2004). ADAR1 and ADAR2 have both been shown to edit pri-miRNA22 transcripts and display overlapping yet distinct site specificities. It has been shown that edited dsRNA is not a good substrate for Dicer (Scadden and Smith, 2001) however, the physiological consequences of this editing are unknown as is the normal function of miRNA22.

1.5 RNA BINDING

RNA binding proteins have been found to be involved in many diverse cellular functions, influencing and regulating RNA molecules. The dsRNA binding motif (dsRBM) was first identified in the *Xenopus laevis* RNA-binding protein A (XlrbpA) and the *Drosophila* maternal effect protein Staufen (St Johnston et al., 1992). Since then a number of other proteins have been identified as containing one or more dsRBMs often in tandemly arranged copies. These include the mammalian dsRNA-dependent protein kinase (PKR), the TAR RNA-binding protein (TRBP), RNase III, Dicer and the members of the ADAR family.

1.5.1 The structure of dsRNA binding motifs

The dsRNA binding motif is approximately 70 amino acids in length and a number of motifs have been analysed by NMR revealing a common α–β–β–β–α structure in which the two α-helices lie on one face of a three-

stranded anti-parallel β-sheet (Bycroft et al., 1995; Kharrat et al., 1995; Nanduri et al., 1998). *In vitro* studies have shown that dsRBDs bind any dsRNA of at least 16bp regardless of its base composition but do not bind to single-stranded RNA or DNA or dsDNA (Bevilacqua and Cech, 1996; St Johnston et al., 1992). A-form dsRNA has its sequence information buried within the major groove (Steitz, 1993) therefore dsRBMs instead rely mainly on non-electrostatic minor groove 2'-OH interactions with only one ionic contact between the protein and dsRNA phosphate backbone (Bevilacqua and Cech, 1996).

A 1.9Å resolution crystal structure of the second dsRBM of XlrbpA complexed with dsRNA confirms the biochemical analysis and shows the motif spanning 16bp of mainly A-form RNA helix and interacting with two successive minor grooves and the intervening major groove of the sugar-phosphate backbone without any direct base contact (Ryter and Schultz, 1998). dsRBD specificity has been well documented. In the case of the ADARs, site-specific editing has been observed *in vitro* indicating no other co-factors are involved. Indeed, ADAR1 and ADAR2 are both capable of binding to the glutamate receptor B subunit transcript yet deaminate different adenosines within it (Melcher et al., 1995). This suggests that any discrimination between specific editing sites by dsRBDs is not mediated by a specific sequence but rather influenced by the secondary and tertiary structural characteristics of the region around the editing site.

The proposal that dsRBMs recognise secondary and tertiary elements of dsRNA was demonstrated by Ramos *et al.* in their NMR study of the third dsRBM of *Drosophila* Staufen. The protein was complexed with an RNA stem loop containing 12 uninterrupted base pairs and protein-RNA interactions were detected along the sugar-phosphate backbone of the nucleic acid. However, additional contacts were observed between the first α-helix of the dsRBM and the single-stranded tetraloop capping the RNA helix (Ramos et

al., 2000). More recently, the specific recognition of a physiologically relevant RNA transcript by the dsRBD of Rnt1p RNase III was found to involve a specific interaction between the fold of a tetraloop structure within the RNA and the two helical turns of the N-terminal helix α1 (Wu et al., 2004).

Due to the highly constrained nature of the tetraloop structure, the recognition of these elements demonstrates how dsRBMs convey specificity for particular RNA targets in a sequence independent manner.

The residues identified in Staufen's third dsRBM that mediate the tetraloop interaction are not conserved across all dsRBMs suggesting that individual dsRBMs, even those within a single protein, are capable of discriminating amongst different RNA molecules. Doyle *et al.* demonstrated that the individual dsRBMs from the *X. laevis* ADAR1 enzyme each recognise different chromosomal sites on transcriptionally active lampbrush chromosomes in a specific manner indicating that each individual dsRBM may recognise distinct and different elements with an RNA molecule (Doyle and Jantsch, 2003).

More recently the structure of the human GluR-B R/G stem loop pre-mRNA has been resolved by NMR (Stefl and Allain, 2005). A stable GCU(A/C)A pentaloop fold was revealed and owing to its similarity to the tetraloop structure described by Wu *et al.* this suggests that the pentaloop may play a role in R/G site recognition (Stefl and Allain, 2005)

1.5.2 dsRNA binding domains can mediate protein dimerisation

As studies progress it has become increasingly apparent that dsRBDs do not solely exist to facilitate the binding of proteins to dsRNA molecules. Instead, as many of these dsRBD-containing proteins have more than one motif, some of which display no dsRNA binding activity, other functions have been identified by biochemical and deletion analysis and assigned to these

domains. One of these alternative functions is the ability of a number of dsRBDs to mediate the dimerisation of their proteins.

A number of dsRBD containing proteins have been found to function as dimers including PKR (Patel et al., 1995), TRBP (Cosentino et al., 1995), RNase III (Lamontagne et al., 2000), and XlrbpA (Hitti et al., 2004).

The most studied of this group is the mammalian dsRNA-dependent protein kinase, PKR. The PKR-PKR interaction was found to be mediated by the dsRNA binding domain and is dsRNA dependent (Cosentino et al., 1995). A K60A point mutation within the dsRBD was found to abolish dsRNA binding activity but have no effect on protein dimerisation indicating that the dsRNA binding and dimerisation activities of the domain are distinct and can be disassociated (Patel et al., 1995). Further studies demonstrate that the converse is also true and dimerisation is not required for dsRNA binding. Instead, dimerisation is mediated by hydrophobic residues located on one side of the amphipathic α-helix of the dsRBD (Patel and Sen, 1998). Indeed, the *in vivo* activation of the kinase domain in PKR is dependent on the binding of dsRNA which in turn mediates dimerisation leading to autophosphorylation and biochemical activation (Patel and Sen, 1998; Ung et al., 2001; Wu and Kaufman, 1997; Zhang et al., 2001).

More recently a member of the ADAR family, the *Drosophila* ADAR protein, has been shown to function as a dimer, mediating this interaction via its N-terminus and first dsRBD (Gallo et al., 2003). Abolition of RNA binding activity in dADAR by point mutations within the first dsRBD also disrupted dimerisation indicating the requirement of dsRNA for dimer formation. However, binding and dimerisation activities can be disassociated by deletion of the N-terminus resulting in a monomer capable of interacting with dsRNA but unable to catalyse the A-to-I editing reaction (Gallo et al., 2003). The dADAR protein exists in a number of different isoforms generated by alternative splicing and self-editing (Palladino et al., 2000a).

These various isoforms have differing editing activities (Palladino et al., 2000a) and can interact with each other via the N-terminus and first dsRBM forming heterodimers (Gallo et al., 2003). This provides another mechanism for the regulation of ADAR editing activity *in vivo*.

1.5.3 Other roles for the dsRNA binding motifs

Eckmann *et al.* identified the third dsRBM of human ADAR1 as a nuclear localisation signal. Deletion analysis indicated that only the entire motif could exhibit NLS activity and that this activity is independent of dsRNA binding suggesting a dual function for this region of the protein (Eckmann et al., 2001).

Studies looking at the different roles of each of the five Staufen dsRBMs demonstrated that neither dsRBM2 nor dsRBM5 were capable of binding dsRNA *in vitro* but played important roles in the microtubule-dependent localisation of oskar mRNA and its translational control (Micklem et al., 2000). This suggests that other dsRBMs not able to bind dsRNA may fulfil other unique roles.

1.6 Substrate specificty of ADARs

1.6.1 Selection via dsRNA Binding domains

dsRNA binding domains are found in a number of proteins with a variety of functions. It is obvious that the role of the dsRBD in each protein differs however, it is unclear how much influence these binding domains have over the catalytic activity.

To ascertain the role of dsRBDs in specific enzymatic function Liu *et al.* constructed a chimeric protein in which the dsRBMs of ADAR1 were substituted for the dsRBMs of PKR (Liu et al., 2000). In chimeras containing

the two PKR dsRBMs and retaining the third dsRBM and catalytic domain of ADAR1, editing activity on GluR-B sites was reduced when compared to wild-type ADAR1 and editing at site A of the 5-HT_{2C} serotonin receptor was abolished. The chimera containing the two PKR dsRBMs plus the ADAR1 catalytic domain was only able to edit a synthetic dsRNA substrate and had no activity on any physiological substrates (Liu and Samuel, 1996). These results suggest an important role for the dsRBD of ADAR1 in the site-specific selection and recognition of target substrates. Especially important is the role and position of the third dsRNA binding motif of ADAR1. Previous studies have found it to be essential for ADAR1 function (Lai et al., 1995) (Liu and Samuel, 1996) and the results from the chimera analysis provide further evidence of its importance.

It should be noted that the activity of the PKR-ADAR1 chimeras was inhibited by known RNA antagonists of PKR indicating that the reduction in ADAR catalytic activity was perhaps not solely due to a loss of function within the dsRBD but rather a gain of function in the chimera in which the PKR motifs have retained their ability to recognise and interact with particular RNA molecules (Liu et al., 2000).

Since the initial identification of a conserved dsRNA binding domain (St Johnston et al., 1992), many dsRBD-containing proteins have been found to have multiple dsRNA binding motifs tandemly arranged. Mutation and deletion analyses have demonstrated that the motifs within a dsRNA binding domain are not functionally equivalent (Lai et al., 1995; McMillan et al., 1995) displaying varying affinities for dsRNA (Bass et al., 1994; Krovat and Jantsch, 1996; Tian and Mathews, 2001) with some possessing no dsRNA binding capabilities at all (Krovat and Jantsch, 1996; Micklem et al., 2000). Indeed, it has been demonstrated that dsRBMs have distinct properties, not necessarily involving dsRNA binding (Micklem et al., 2000), but that some enzyme functions rely on cooperation between two or more motifs (Tian and Mathews, 2001).

It has been shown that the dsRBMs of *Xenopus* ADAR1 are responsible for chromosomal targeting and that the individual motifs are capable of binding to different sites in a distinct and specific manner (Doyle and Jantsch, 2003). This indicates that each motif may play a role in the selection of specific editing sites.

More recently, Stephens *et al.* demonstrated that the dsRBMs of ADAR2 bind selectively on their own to particular contact points adjacent to the GluR-B editing site on a model RNA duplex (Stephens et al., 2004). The dsRBMs of PKR were also able to bind to this model RNA duplex but displayed different binding preferences. Benzyl modification of guanosine 2-amino groups blocks the ADAR2 dsRBM recognition surfaces resulting in a decrease in deamination rate thus demonstrating a correlation between ADAR2 deamination activity and dsRBM binding selectivity (Stephens et al., 2004).

Due to the structural characteristics of an A-form dsRNA helix it is apparent that dsRBDs do not bind in a sequence–specific manner instead it appears they rely on the recognition of secondary and tertiary characteristics of the dsRNA molecule. A number of dsRNA binding proteins have been shown to specifically interact with stem-loop structures within an RNA molecule (Ramos et al., 2000; Spanggord and Beal, 2001; Wu et al., 2004). This ability to recognise a particular structural element of RNA allows for the identification of target molecules.

Internal loops, mismatches and bulges within dsRNA substrates edited by ADARs have been shown to increase selectivity (Bass et al., 1994; Nishikura et al., 1991). It has been shown that these internal loops convert long promiscuously edited dsRNA duplexes into a series of short, more selectively edited substrates (Lehmann and Bass, 1999). This is an example of how RNA structure and therefore its interaction with the dsRBD can influence the efficiency and activity of the ADARs.

As well as the binding capability and functional characteristics of the individual dsRBMs it has been suggested that the linker region between these motifs can influence the activity of the enzyme. NMR studies have shown the dsRNA binding domain of PKR to consist of two dsRBMs flanking a 22 residue highly flexible linker (Nanduri et al., 1998). This linker region may enable the dsRBMs to wrap around the dsRNA duplex and bind co-operatively.

ADAR1 and dADAR have both been shown to have alternative splice sites located in linker regions between dsRBMs (Liu et al., 1997) (Palladino et al., 2000a). The presence or absence of these alternative exons alters the spacing between the dsRBMs potentially altering the binding characteristics of the entire domain. Palladino *et al.* demonstrated that the absence of the alternatively spliced exon 3a in *Drosophila* ADAR creates a shorter linker region between the two dsRBMs and results in a more active isoform of the ADAR enzyme (Palladino et al., 2000a).

1.6.2 Selection via the deaminase domain

Though they are able to recognise specific pre-mRNA that undergo editing, the dsRBMs do not appear to have the capacity to select the specific adenosine to be deaminated. This specificity is mediated via the catalytic domain. When the deaminase domains of ADAR1 and ADAR2 are exchanged, the resulting chimeras editing site preferences reflect the origin of the deaminase domain indicating that it is this region of the ADAR enzymes that plays a defining role in substrate specificity (Wong et al., 2001).

A subset of the ADAR family of enzymes, the ADATs, have no known nucleic acid binding domain yet target specific adenosines on target tRNAs effectively (Gerber and Keller, 2001). This begs the question, are the dsRBDs essential for ADAR function? Herbert *et al.* constructed a deletion mutant of ADAR1 consisting of only the deaminase domain and in a transfection assay

this mutant successfully edited a 15bp dsRNA stem-loop structure (Herbert and Rich, 2001). More recently, Macbeth *et al.* demonstrated that the deaminase domain of hADAR2 alone was able to edit a non-specific dsRNA substrate *in vitro* (Macbeth et al., 2005). In addition, studies on the hepatitis delta virus amber/W site have shown that ADAR1 is able to edit a hairpin consisting of the amber/W site with 4bps upstream and 4bps downstream of the editing site (Sato et al., 2001). This is the smallest substrate described and is likely too small a target for the dsRBD to bind. Taken together these results suggest that the dsRBD is not essential for editing but instead may play a role in increasing the efficiency of the editing reaction on more complex substrates.

Sequence analyses around editing sites have shown that the editing efficiency of ADAR1 and ADAR2 varies depending on the nucleotide opposite the editing site. ADARs are more efficient when there is an A:C mismatch at the editing site than an A:U base pair or A:G/A:A mismatch (Stephens et al., 2000; Wong et al., 2001). ADAR1 has also been shown to have a 5′ neighbour preference, favouring adenosines with a 5′ neighbour A = U > C > G (Bass et al., 1994). ADAR2 also exhibits a 5′ neighbour preference of A ≈ U > C = G however, sequence analysis has also identified a 3′ neighbour preference for ADAR2 of U = G > C = A (Lehmann and Bass, 2000).

These data suggest that as well as the secondary and tertiary structural characteristics of the dsRNA substrate recognised by the dsRNA binding motifs there is an element of enzyme specificity that is influenced by the catalytic domain.

1.7 CATALYTIC MECHANISM OF ADENOSINE DEAMINATION

1.7.1 The role of the Deaminase domain

The catalytic domains of the ADAR enzymes contain the conserved residues, HAE$(x)_{41-58}$PCG$(x)_{44-154}$SCSDK, analogous to sequences in the active sites of the cytidine deaminase family of enzymes. Based on studies of two nucleoside deaminase members of this family, adenosine deaminase (ADA) and cytidine deaminase (CDA) (Carter, 1998), the cysteine and histidine residues are the ligands to the active site zinc, while the glutamate residue is necessary for proton transfer function. Site-directed mutagenesis studies of ADAR1 demonstrated that each of these residues was required for enzyme function (Lai et al., 1995).

The cytidine deaminases are metalloenzymes that use a zinc-activated water molecule for the hydrolytic deamination of their nucleoside substrates. Polson *et al.* demonstrated that in ADAR reactions the source of the oxygen atom incorporated at C6 of the inosine product is water suggesting a nucleophilic attack at C6 of the adenosine nucleotide (Polson et al., 1991). More recently, the 1.7Å resolution crystal structure of the catalytic domain of hADAR2 revealed a zinc ion within the active site and confirmed the role of the histidine and cysteine residues in co-ordinating zinc and the hydrogen bonding of a water molecule by the glutamate (Macbeth et al., 2005). Taken together, this data would suggest a common evolutionary origin for the catalytic domain of ADARs and members of the cytidine deaminase family.

Whilst the reaction mechanism of the ADARs may be similar to that of the nucleoside deaminases, the target substrates vary greatly. ADARs deaminate adenosine nucleotides within a dsRNA duplex by a nucleophilic attack at C6 of the adenosine nucleotide. This C6 atom is buried deep within

the major groove of the A-form helix therefore the ADAR enzyme may be required to flip its target nucleotide out of the RNA helix core. By studying enzyme-induced changes in the fluorescence of 2-AP-modified substrates, Stephens *et al.* demonstrated that, like the adenosine methyltransferases, ADAR2 appears to flip out the reactive adenosine into a binding pocket (Stephens et al., 2000) and that this mechanism is dependent on the presence of the catalytic domain of ADAR2 (Yi-Brunozzi et al., 2001).

Interestingly, a molecule of inositol hexakisphosphate (IP_6) was also found within the core of the catalytic domain and shown to interact with many arginine and lysine residues (Macbeth et al., 2005). The *S. cerevisiae* mutant *ipk1*Δ is unable to synthesis IP_6 and expression of hADAR2 and ADAT1 in these mutants resulted in inactive proteins indicating that IP_6 is required for ADAR activity (Macbeth et al., 2005).

1.7.2 The role of the dsRNA binding domain

On binding a dsRNA substrate, ADAR2 has been shown to induce a number of conformational changes. The footprint of the dsRNA binding domain protects a region of duplex surrounding the editing site and on binding, nucleotides on the non-edited strand become hypersensitive to hydrolytic cleavage indicating an increase in the conformational flexibility of this region (Yi-Brunozzi et al., 2001). For many base-flipping enzymes, the separation of the phosphodiester backbone is proposed to lower the activation energy barrier required for base flipping. This increase in flexibility initiated by dsRBD binding may act to lower the energy barriers to base flipping during the ADAR reaction. It should be noted that binding of the dsRBD of PKR to the same dsRNA molecule did not induce this conformational change in dsRNA structure indicating that this is a function particular to the ADAR dsRNA binding domains (Yi-Brunozzi et al., 2001).

1.8 Regulation of ADARs

1.8.1 The co-ordination of RNA editing and splicing

The A-to-I editing of a number of physiological substrates must occur before or simultaneously to splicing as the dsRNA structure required for editing is formed between exonic and downstream complementary intronic sequences (Burns et al., 1997; Herb et al., 1996; Higuchi et al., 1993; Lomeli et al., 1994). This suggests a possible regulatory role for splicing in the dsRNA editing process. Indeed, the association of enzymatically active ADAR1 and ADAR2 proteins with spliceosomal Sm and SR proteins within the large nuclear ribonucleoprotein (lnRNP) would suggest a link between splicing and RNA editing (Raitskin et al., 2001).

Bratt *et al.* demonstrated that *in vitro* editing and splicing of the glutamate receptor pre-mRNA inhibit each other (Bratt and Ohman, 2003). Exposing the transcript to splicing factors prior to ADAR2 abolished editing whilst the presence of ADAR2 appears to sequester the 5′ splice site required for splicing to occur. This inhibition of splicing was overcome by the addition of RNA helicase A suggesting it out competes ADAR2 in binding to this transcript (Bratt and Ohman, 2003). Further studies *in vivo* indicate that editing and splicing appear to happen sequentially during RNA pol II transcription suggesting a regulatory role for the transcription machinery in co-ordinating these RNA processing events (Bratt and Ohman, 2003).

1.8.2 Other factors potentially involved in ADAR regulation

As the ADAR3 enzyme currently has no known physiological substrates and demonstrates no *in vitro* editing activity, it has been suggested that it plays a role in regulating the activity of the other members of the ADAR family. It has been shown to inhibit the *in vitro* editing activities of both ADAR1 and ADAR2 though the mechanism for this is not understood (Chen et al., 2000).

Likely, it is the result of competitive binding to the dsRNA substrate as ADAR3 is capable of binding to dsRNA (Chen et al., 2000) and further studies have suggested that ADAR3 is unable to form heterodimers with ADAR1 and 2 (Cho et al., 2003).

Liu *et al.* demonstrated that the Vaccinia Virus E3L interferon resistance protein is a potent inhibitor of ADAR1 enzymatic activity on synthetic dsRNA substrates (Liu et al., 2001). E3L also weakly inhibited the site-selective editing activity by ADAR1 at the R/G site of the glutamate receptor B subunit pre-mRNA and the A site of serotonin 2C receptor pre-mRNA. Deletion analysis of E3L protein indicated that the carboxy-proximal dsRBM was essential for this antagonistic function (Liu et al., 2001). The precise physiological function of this anti-ADAR1 activity is unclear but E3L is known to antagonise other interferon-inducible proteins (Romano et al., 1998) and so it is likely a part of the overall antagonism of interferon mediated by the E3L viral protein.

Small nucleolar (sno) RNAs have been shown to modify nucleotides within rRNAs and spliceosomal snRNAs. There are 2 classes of guide snoRNAs; the fibrillarin-associated box C/D snoRNAs responsible for the methylation of the ribose moiety at the 2'-hydroxyl group and the box H/ACA snoRNAs responsible for the conversion of uridines to pseudouridine (Kiss, 2001).

With the identification of a number of novel snoRNAs in the brain tissue of humans and mice, the possibility that these molecules may also modify mRNAs has been hypothesised (Cavaille et al., 2000; de los Santos et al., 2000). In particular, the identification of a novel box C/D snoRNA, HBII-52, which carries an 18 nucleotide region perfectly complementary to the 5-HT_{2C} mRNA suggests that modification of mRNAs may be possible (Cavaille et al., 2000). Furthermore, the target nucleotide for 2'-*O*-methylation is the adenosine at editing site C of the serotonin receptor (Cavaille et al., 2000). As methylation at the 2'-OH group results in a decrease in deamination

efficiency (Yi-Brunozzi et al., 1999), the possibility that this C/D snoRNA may inhibit editing exists. More recently it has been demonstrated that MBII-52 is indeed a C/D snoRNA and specifically decreases ADAR2-mediated nucleolar editing at site C of the serotonin 5-HT_{2C} receptor mRNA (Vitali et al., 2005).

1.9 Aims

The aims of this work are to investigate the properties of the individual dADAR functional domains.

Many studies have indicated that dsRNA binding motifs often have distinctive roles that do not always involve binding to dsRNA (Cosentino et al., 1995; Eckmann et al., 2001; Gallo et al., 2003; Micklem et al., 2000). The *Drosophila* ADAR protein has 2 dsRNA binding motifs separated by a linker region containing an alternatively spliced exon, 3a, whose presence or absence can alter the editing activity of full length dADAR (Palladino et al., 2000a). Structural analysis of these motifs will determine whether they are *bona fide* members of the dsRBM family and by assaying their binding capacity on dsRNA we may gain an insight into the relative functions of these motifs. Indeed, it has already been shown that dsRBM1 plays a role in dADAR protein dimerisation (Gallo et al., 2003).

The *Drosophila* ADAR protein is a member of the cytidine deaminase (CDA) family (Palladino et al., 2000a). Other CDA family members, such as the ADATs, do not contain any known RNA binding motifs yet are capable of editing highly structured RNA molecules, such as tRNA, accurately and efficiently. This begs the question; can the deaminase domain of an ADAR protein function without the dsRNA binding motifs? Expression of the dADAR deaminase domain and assay of its editing function will determine whether the dsRNA binding motifs are required for editing activity and perhaps further elucidate their role also.

A number of studies have attempted to clarify how the ADARs specify which adenosine within a dsRNA molecule will be edited. Much of this specificity in the human proteins has been assigned to the deaminase domain (Wong et al., 2001) though more recently the dsRNA binding domain has also been shown to play a role (Stefl et al., 2006). By comparing the editing patterns of the dADAR deaminase domain and the full length protein on dsRNA it will be possible to determine whether editing site specificity is also conferred by the deaminase domain of dADAR.

The binding of RBMs to dsRNA has been shown to induce a conformational change in a dsRNA substrate (Yi-Brunozzi et al., 2001) and therefore may facilitate the editing activity of the deaminase domain. The possibility of an interaction between these individually expressed domains maybe directly, but more likely via a substrate dsRNA, may demonstrate how these domains function together within the full length ADAR enzyme.

2

MATERIALS AND METHODS

All reagents were obtained from Sigma or Roche unless otherwise stated.

2.1 MOLECULAR BIOLOGY

2.1.1 Bacterial Strains

Competent XL1-blue or commercially produced DH5-α (subcloning efficiency, Invitrogen) cells were used for all cloning procedures. BL21 (DE3) cells were used for protein expression. Cultures were grown in Luria-Bertani (L-broth) medium (1% (w/v) tryptone, 0.5% yeast extract, 1% NaCl) or on L-agar plates (1% tryptone, 0.5% yeast extract, 1% NaCl, 1.5% agar) at 37 °C.

2.1.2 Generating competent bacteria

1 ml of an overnight culture of *XL1*-blue or *BL21* was grown to an absorbance at 600 nm of 0.5 - 0.6 (using the sample equipment as described in section 2.1.6.5). Cells were pelleted at 4500 x *g* for 5 min at 4°C and resuspended in 0.4 original volume of ice cold TFB1 (30 mM $C_2H_3KO_2$, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM RbCl, 15% glycerol). The resuspended cells were incubated on ice for 5 min at 4°C before being centrifuged again as before. The pelleted cells were then resuspended in 1/25 original volume of ice cold TFB2 (100 mM MOPS or PIPES, pH 6.5, 75 mM $CaCl_2$, 10 mM RbCl, 15% (v/v) glycerol) and incubated on ice for 15 – 60 min. Aliquots were snap frozen on dry ice and stored at –70 °C until required.

Electro-competent cells for high efficiency transformation were also prepared. 1 ml of an overnight culture of XL1-blue or BL21 was grown at 37 °C to an absorbance at 600 nm of 0.5 - 1.0. The flask was chilled on ice for 30 min and the cells harvested by centrifuge at 4,000 x *g* for 15 min at 4 °C. The pellets were resuspended in 1 litre of cold water and centrifuged as before. The pellets were then washed in 0.5 litres cold water and centrifuged once more. Finally the pellets were washed in 20 ml of 10% glycerol, centrifuged and the pellet was resuspended in a final volume of 2 to 3 ml of 10% glycerol. The suspension was snap frozen in dry ice in 40 µl aliquots and stored at -70 °C.

2.1.3 Bacterial Transformation

Competent cells were incubated on ice with plasmid DNA or ligation product (Section 2.1.6.7) for at least 15 min. Transformations were then heat shocked at 42 °C for 1½ min. After addition of 1 ml of L-Broth, reactions were incubated for at least 45 min at 37 °C before being plated on L-agar plates containing the appropriate antibiotic selection (Ampicillin at 50 µg/ml, Kanamycin at 50 µg/ml). Transformation plates were then grown overnight at 37 °C.

Alternatively for higher efficiency the plasmid DNA or ligation product could be electro-transformed. 1-2 µl of plasmid DNA was mixed with 40 µl of competent *E. coli* cells in an ice-cold 0.2 cm electroporation cuvette and electroporated at 2.5 kV, 25 µF, 200 Ω (Bio-Rad Gene Pulser). Immediately 1 ml of ice-cold SOC medium (2% Bacto trypyone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) was added to the cuvette and the cells quickly resuspended. The cell suspension was transferred to a 15 ml sterile tube and incubated at 37 °C for 1 h before being plated on L-agar plates containing the appropriate antibiotic selection (Ampicillin at 50 µg/ml, Kanamycin at 50 µg/ml). Transformation plates were then grown overnight at 37 °C.

2.1.4 Isolation of plasmid DNA

5 or 50 ml cultures of single bacterial colonies were grown overnight in L-Broth at 37 °C. Plasmid preparations were then harvested from these cultures using commercial mini- or midiprep kits (Qiagen). Briefly, bacteria are lysed under alkaline conditions. The lysate is then neutralised and passed through a silica-gel membrane that selectively absorbs DNA. The membrane is then washed before the plasmid DNA eluted.

2.1.5 Purification of DNA

2.1.5.1 Phenol/Chloroform extraction

An equal volume of buffered phenol/chloroform (50% phenol/48% chloroform, 2% isoamylalcohol, pH 8.0) was added to the DNA preparation and agitated. This was centrifuged at ~12,000 x *g* for 1 minute. The top (aqueous) phase was transferred to a fresh tube, avoiding any white precipitate that may be present at the boundary between the two phases and an equal volume of chloroform added. Agitation and centrifugation was repeated. Again the top layer was removed. Ethanol precipitation was then performed to concentrate the DNA (section 2.1.5.2).

2.1.5.2 Ethanol precipitation

$1/10^{th}$ volume 5 M NaCl and 3 volumes of ice cold 100% ethanol were added to the DNA sample, the tube was vortexed and then centrifuged at ~12,000 x *g* for 30 min at 4 °C. The supernatant was discarded and the pellet washed in ice cold 70 % (v/v) ethanol with dH_2O. Centrifugation was repeated for 10 min and the supernatant discarded. The pellet was dried at room temperature and resuspended in an appropriate volume of dH_2O.

2.1.6 Analysis and generation of recombinant DNA

2.1.6.1 Plasmid DNA digestion

Plasmid DNA was digested with restriction enzymes (New England Biolabs) in reaction buffer in accordance with the manufacturers' instructions. Digestion products were then analysed and/or purified by gel electrophoresis (section 2.1.6.3) and used for further cloning steps (section 2.1.6.7).

2.1.6.2 PCR amplification of DNA fragments

Standard conditions for a 50 µl reaction were : 50 ng template DNA, 5 µl 10x PCR buffer containing 15 mM $MgCl_2$ (Roche), 0.2 mM Ultrapure dNTPs (Amersham), 1 µM each forward and reverse primers, 2.5 Units Taq Polymerase (Roche).

The standard PCR program used was a 94 °C hot start before 35 cycles of melting at 94 °C for 30 seconds, annealing at 55 °C for 30 seconds and elongation at 72 °C for 30 seconds. A final elongation step at 72 °C was carried out for 8 min.

DNA fragments for cloning were generated by *Pfu*Turbo DNA Polymerase (Stratagene) amplification. This proofreading enzyme was chosen for its increased yield and lower error rate when amplifying products. The reaction mix was as follows: 100 ng DNA template, 100 ng of forward and reverse primers, dNTP, polymerase and buffer diluted to manufacturer's recommendations in dH_2O. The amplification conditions were as follows: a 95 °C hot start, 30 cycles of 95 °C for 30 seconds, 68 °C for 30 seconds, 72 °C for 60 seconds and a final extension of 10 min at 72 °C. Primers are listed in Table 2.1.

Table 2.1 PCR Primers

RBM FW	AAT AGC GCT AGC AAG AAC ACG GTG GCC ATG CTG
RBM REV	CTA CTG GGA TCC TTA AAT ATT GCA TAA CGA AGC AAG
RBM1 REV	CTA CTG GGA TCC TTA CTT AAA CTG TAT AAA GCT GCG
RBM2 FW	CAG TCA GCT AGC GAT AAG GGT CCT GTC ATG CTC
DA REV	TGA CTG GGT ACC TTC GGC AAG ACC GAA CTC GTC
DA FW	CAG TCA GAA TTC GAC GAC AAG TCG TCA TCG
SP6 Promoter	ATT TAG GTG ACA CTA TAG AA
T7 Promoter	TAA TAC GAC TCA CTA TAG GG
KP 5' T7 antisense	GAG ACC GGA ATT CGG ATC
KP 3' SP6 sense	GAT CCT CYA GAG YCG AYC
KP 5' SP6 sense	GAA TAC AAG CTT GTC GRG

2.1.6.3 Site-directed mutagenesis

Point mutations were generated by oligonucleotide-directed mutagenesis using the QuikChange™ site-directed mutagenesis kit (Stratagene). The reaction mix was as follows: 50 ng DNA template, 100 ng of forward and reverse primers, dNTP, polymerase and buffer diluted to manufacturer's recommendations in dH_2O. The amplification conditions were as follows: a 30 seconds 95 °C hot start, 12 cycles of 95 °C for 30 seconds, 55 °C for 1 min,

and extension at 68 °C for 1 min / kb of template plasmid. In the case of human *adar2* in pPICZa the extension time was 6 min. The primers used in this reaction are listed in table 2.2.

Table 2.2 Primers for oligonucleotide-directed mutagenesis

adar2 S/G FW	GCC AGA ATC TTC GGA CCA CAT GAG CCA ATC
adar2 S/G REV	GAT TGG CTC ATG TGG TCC GAA GAT TCT GGC

1 µl *Dpn*I was added to each PCR reaction and incubated for 1 h to digest the parental methylated DNA. The PCR reaction was then cleaned and 1 µl transformed into electo-competent XL1-blue cells.

2.1.6.4 Agarose Gel electrophoresis

1% horizontal agarose gels in 0.5x TBE (45 mM Tris-HCl, 45 mM boric acid, 0.05 mM EDTA pH 8.0) were used with 0.5x TBE electrophoresis buffer to resolve DNA samples, PCR products and plasmid digestion products. All samples had 1 volume of 6x loading buffer (15% Ficoll 400, 0.25% Bromophenol Blue, 0.25% xylene cyanol) added to 5 volumes of sample. Commercially available DNA size markers (100 bp and 1 Kb ladder Promega and Invitrogen, respectively) were diluted to 50 ng/µl and 500 ng loaded per well to enable size determination and quantification of DNA fragments.

Ethidium bromide (2,7-diamino-10-ethyl-9-phenyl-phenathridium bromide), a DNA intercalating stain, was added to molten agarose before pouring a gel. Stained DNA was then visualised with ultraviolet light from a transilluminator.

2.1.6.5 Gel purification of DNA fragments

For purification, agarose slices containing the appropriate DNA fragments were dissected from gels using clean razor blades. The DNA was extracted

from the agarose using a commercial kit (Qiagen). Briefly, gel slices are melted in equilibration buffer. The solution is then passed over a silica-gel membrane that selectively binds DNA. Washing this membrane removes ethidium bromide and residual agarose. The DNA was then eluted in the required volume of dH_2O.

2.1.6.6 Measuring quality and quantity of DNA

DNA concentration was measured spectrophotometrically and/or by gel electrophoresis alongside DNA of known concentration. DNA was diluted 1:100 in TE and transferred to a quartz cuvette. The absorbance 260 nm (A_{260}) was measured for DNA. An A_{260} of 1 corresponds to ~50 µg/ml of DNA. To determine the purity of nucleic acid, the A_{280} was also measured. Pure DNA has an A_{260} /A_{280} of ≥ 1.8 and values lower than these indicate contamination by proteins, RNA or phenol.

2.1.6.7 Dephosphorylation of DNA fragments

Linear plasmid DNA digested with the appropriate enzymes for a cloning step was treated with alkaline phosphatase in accordance with the manufacturers' instructions. This removes 5′ phosphates from the ends of the DNA backbone preventing reformation of empty circular plasmid DNA. This subsequently increases the yield of positive ligation products as successful ligation should only be achieved when phosphates are donated to the backbone ends by untreated insert.

2.1.6.8 Ligation of DNA fragments

50 ng of digested vector was incubated with 3 times the molar ratio of cloning insert. 400 U (1 µl) T4 DNA ligase enzyme (New England BioLabs) and the appropriate volume of 10x reaction buffer were used per ligation reaction, typically in a total volume of 10 µl. The reactions were incubated at

room temperature for one hour or overnight at 16 °C. Subsequently, 5 µl of a reaction were then used for bacterial transformation (section 2.1.3).

2.1.6.9 Sequencing of DNA fragments

All cloning products were sequenced to monitor the integrity of recombinant plasmids and inserts. In a protocol similar to Sanger di-deoxy sequencing, BigDye Terminator v2.0 (ABI Prism) chemistry was used as follows: 200 ng of linear plasmid DNA, 0.8 µM sequencing primer and BigDye were diluted in dH_2O in accordance with the manufacturers' instructions.

Cycle conditions used for sequencing reactions were 25 cycles of; melting at 96 °C for 30 seconds, annealing at 50 °C for 15 seconds followed by elongation at 60 °C for 4 min. Sequencing reactions were precipitated with 2 µl of 3 M Sodium Acetate (pH 4.6) and 50 µl of 100% ethanol at room temperature and out of bright light for 20 min. Samples were then centrifuged at ~12,000 x *g* for 20 min at 18 °C and the supernatant was removed. Pellets (not always visible) were washed with 150 µl of 70% ethanol and air dried.

2.1.7 Isolation of Genomic DNA from *Pichia pastoris*

5 ml of BMGY was inoculated with a single yeast colony. At stationary phase the culture was pelleted at 4000 x *g* for 5 min. The pellet was resuspended in 250 µl of SP1 buffer (1.2 M sorbitol, 50 mM sodium citrate, 50 mM Na_3PO_4, 40 mM EDTA, pH 5.6) containing 0.4 mg/ml zymolase and incubated for 45 min at 37 °C. The cells were pelleted at 6000 x *g* for 15 seconds and resuspended in 0.5 ml TE and 50 µl 10% SDS before being vortexed. 165 µl of 5 M potassium acetate was added and the pellets were stored on ice for 30 min before being centrifuged for 10 min at 4 °C. The supernatant was added to 750 µl of isopropanol, placed on dry ice for 5 min and then centrifuged for 10 min. The pellet was resuspended in 300 µl of TE

and RNase added to a final concentration of 10 μg/ml. The sample was incubated for 30 min at 37 °C before phenol/chloroform extraction (section 2.1.5.1) and ethanol precipitation (section 2.1.5.2). The genomic DNA was then resuspended in 20 μl of TE.

2.1.8 Isolation of total RNA from *Pichia pastoris*

A 50 ml culture inoculated with a single yeast colony was grown up and protein expression induced as described in section 2.2.1.5. 10 ml of culture was pelleted and resuspended in 500 μl of TE. This was transferred to a microfuge tube and pelleted again before being resuspended in 300 μl of extraction buffer (50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 100 mM NaCl, 1% SDS). Approximately 300 μl of glass beads and 300 μl of phenol were added to the tube and the cells vortexed for 15 min at high speed at +4 °C. The samples were centrifuged for 5 min at 12,000 x *g* and the supernatant transferred to a fresh tube. The sample was phenol/chloroform extracted, ethanol precipitated and the pellet air dryed before being resuspended in 25 μl of dH_2O. The sample was stored at -20 °C.

2.1.9 RT-PCR of yeast mRNA

The yeast RNA was treated with DNase to remove any contamination. 2.5 μl of a 1 μg/μl RNA stock was incubated with 1 μl of 10x DNase buffer, 1 μl of DNase and 5.5 μl dH_2O for 1 hour at 37 °C. 1 μl of 25 mM EDTA was added and the reaction mix incubated for a further 10 min at 65 °C. 12 μl of dH_2O and 2 μl of 0.5 μg/μl oligo dT was added to each RNA sample and incubated at 70 °C for 10 min to allow annealing of the oligo dT to the poly(A) tails of the mRNA molecules. The samples were placed on ice and spun briefly at 4 °C. 4 μl of 10x PCR buffer, 4 μl 25 mM $MgCl_2$, 2 μl of 2.5 mM dNTPs and 4 μl of DTT were added and the samples mixed well and split into two 19 μl aliquots. Each aliquot was incubated at 42 °C for 5 min and then 1 μl of Superscript reverse transcriptase (Gibco BRL) added to one of each pair of

tubes. The sample was incubated for a further 50 min at 42 °C and then 70 °C for 15 min. The sample was placed on ice and 1 μl of each used as template for a 20 μl PCR reaction (section 2.1.6.2).

2.1.10 pGEM®- T Easy cloning

Prior to sequencing PCR products were cloned into the pGEM®-T Easy vector (Promega). The single 3'-T overhangs at the vector insertion site allow for high efficiency cloning of PCR products. 1 μl of PCR product was added to a ligation mix containing buffer, ligase and vector in accordance with the manufacturer's instructions. 5 μl of each ligation reaction was transformed into DH5α cells and plated on L-agar plates containing 50 μg/ml ampicillin, 0.5 mM IPTG and 80 μg/ml X-gal. Plates were incubated overnight at 37 °C and screened the following day for white insert-containing colonies.

2.2 Synthesis and Analysis of ADAR Proteins

2.2.1 Expression and purification of recombinant ADAR proteins in *Pichia pastoris*

2.2.1.1 pPICZ A expression constructs

Drosophila ADAR domains and hADAR2 were amplified (section 2.1.6.2) from full length cDNA and ligated (section 2.1.6.7) into pPICZ A (Invitrogen). The myc and 6 X HIS epitope tag sequences of pPICZ A are expressed in the resulting fusion protein. pPICZ A was also used to construct a derivative called pPICZ A-FLIS6 that contains a short open reading frame encoding a FLAG epitope tag at the N-terminus and a 6 X HIS epitope tag at the carboxy terminus inserted at the *Eco*RI site in the pPICZ A multiple cloning site. A *Spe*I site between the two epitope tags was used to clone ADAR coding sequences in frame to the epitope tags. The myc and 6 X

HIS epitope tag sequences present in the original pPICZ A vector are not expressed in these constructs. For dual expression of 2 domains of the *Drosophila* ADAR protein, the pHIL D2 (Invitrogen) cloning vector was also used.

The ligation reactions were transformed into *E. coli* DH5α cells, plated on low salt LB agar (1% tryptone, 0.5% yeast extract, 0.5% NaCl, 1.5% agar) containing 25 μg/ml Zeocin™ (Invitrogen) and incubated overnight at 37 °C. Single colonies were selected and grown overnight in low salt LB media (1% tryptone, 0.5% yeast extract, 0.5% NaCl) with 50 μg/ml Zeocin™. Plasmid DNA was isolated and checked for inserts by restriction enzyme digestion. Positive clones were sequenced to confirm the presence of the correct open reading frame.

2.2.1.2 Generating competent *Pichia pastoris*

5 ml of YPD (1% Yeast extract, 2% peptone, 2% dextrose) was inoculated with the strains KM71 or KM71H and grown overnight in a 50 ml conical flask at 30 °C. The following day 500 ml of fresh YPD was inoculated with 1 ml of the overnight culture and grown in a 2 litre flask to an OD_{600} of 1.3 – 1.5. The cells are harvested by centrifugation at 1500 x *g* for 5 min at 4°C then washed twice with 500 ml ice-cold sterile water and once with 20 ml ice-cold 1 M sorbitol. Finally, the pellet was resuspended in 1 ml 1 M sorbitol to give a final volume of approximately 1.5 ml. The cells were stored on ice until required.

2.2.1.3 *Pichia pastoris* transformation

5-10 μg of the pPICZ A or pHIL D2 expression constructs was digested with *Sac*I and linearisation checked by electrophoresing a small aliquot on a 1% agarose gel. The linearised DNA was extracted with phenol/chloroform (section 2.1.5.1) and ethanol precipitated (section 2.1.5.2). The pellet was resuspended in 10 μl dH_2O. The linearised DNA was mixed with 80 μl of

competent *Pichia* cells in an ice-cold 0.2 cm electroporation cuvette and electroporated at 1.5 kV, 25 μF, 200 Ω (Bio-Rad Gene Pulser). Immediately 1 ml of ice-cold 1 M sorbitol was added to the cuvette, transferred to a 15 ml sterile tube and incubated without shaking at 30 °C for 1-2 hours. 25 and 100 μl was plated on separate YPDS (1% yeast extract, 2% peptone, 2% dextrose, 1 M sorbitol, 2% Agar) plates containing 100 μg/ml Zeocin™and incubated for 2-5 days at 30 °C until colonies formed.

2.2.1.4 Analysis of *Pichia* transformants

10-20 colonies from transformation plates were picked and streaked for single colonies on fresh YPDS agar plates with 100, 200, 500 and 1000 μg/ml Zeocin™. Plates were incubated for 2-3 days at 30 °C. The transformant lines showing the highest resistance to Zeocin™have the highest copy number of the expression construct.

2.2.1.5 Expression of recombinant ADAR in *Pichia pastoris*

5 ml of buffered glycerol-complex medium (BMGY: 1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1.34% YNB, 4 x 10^{-5}% biotin, 1% glycerol) was inoculated with a *Pichia* colony and grown overnight at 30 °C in a loosely capped 50 ml falcon tube. 500 ml of fresh BMGY was inoculated with the 5 ml starter culture and grown overnight at 30 °C to an OD_{600} of 2-6. The cells were pelleted by centrifugation at 1500 x g for 5 min at +4 °C and resuspended in buffered methanol-complex medium (BMMY: 1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1.34% YNB, 4 x 10^{-5}% biotin, 0.5% methanol) to induce expression. The culture was returned to the original flask and grown overnight at 30 °C. The following day methanol was added to a final concentration of 1% to counter any losses due to evaporation and grown overnight as before.

2.2.1.6 Harvest of *Pichia pastoris*

The cells were pelleted by centrifugation at 1500 x *g* for 5 min at +4 °C and the BMMY supernatant removed. The pellets were washed with 250 ml of autoclaved ice-cold water to remove any remaining BMMY and spun as before. Each pellet was resuspended in 40 ml of Buffer Q/200 (200 mM KCl, 50 mM Tris-HCl, pH 7.9, 20% glycerol, 0.5 mM PMSF, 0.7 mg/ml Pepstatin, and 0.4 mg/ml Leupeptin), divided into 2 x 50 ml falcon tubes and spun as before. At this point the cells were either broken and protein extracted or the pellets were frozen in liquid nitrogen and stored at –70 °C for up to 1 year.

2.2.1.7 Protein extraction

Cells were resuspended in 1 ml buffer Q/200 per gram of cell pellet. The suspension was lysed with 2–3 passes using a French press (Thermo Electron).

2.2.1.8 Ni^{2+}-NTA affinity chromatography of *Drosophila* ADAR and hADAR2

Lysed *Pichia* cells were centrifuged for 15 min, at 20,000 x *g* at +4°C to separate cell debris. The supernatant was removed and combined with Ni^{2+}-NTA resin (Qiagen) previously equilibrated with load buffer (buffer Q/200 + 10 mM imidazole, 0.5% Triton X-100) and rotated gently for 1-2 hours at 4 °C. The resin mix was loaded onto a 20 ml Econo-column (Bio Rad) 3 times at 4 °C and the flow through collected. The column was washed with 2 x 10 ml of wash buffer (Buffer Q/200 + 20 mM Imidazole, 0.5% Triton X-100) and fractions collected. The protein was eluted with elution buffer (Buffer Q/200 + 250 mM imidazole, 0.5% Triton X-100) in 10 x 500 µl fractions. At each stage all fractions are collected and stored in –70 °C. 30 µl aliquots of crude extract, flow through, wash and each eluate were collected for analysis by SDS-PAGE (section 2.2.3.1).

2.2.1.9 Anti-FLAG affinity column purification of recombinant ADARs

It is possible to further purify the Ni^{2+}-NTA eluates using the N-terminal FLAG epitope tag. The eluates are loaded on to an anti-FLAG M2 affinity gel in which a purified murine monoclonal antibody is coupled to agarose. The protein is eluted with a FLAG peptide. A mini-column (0.7 cm diameter) was poured containing a 150 μl bed volume of anti-FLAG M2 matrix and was equilibrated with buffer Q/200. The column was washed by loading three sequential 5 ml aliquots of 0.1 M glycine HCl, pH 3.5 followed by three sequential 5 ml aliquots of buffer Q/200 to activate the resin. The fractions containing dADAR from the Ni^{2+}-NTA column were pooled and loaded on to the anti-FLAG M2 affinity column. The flow through was collected and passed over a further 2 times to maximise binding. The column was washed three times with 12 ml of buffer Q/200 and the protein was eluted from the column with 5 X 500 μl fractions of elution buffer (buffer Q/200 plus FLAG peptide 100 μg/ml). At each stage all fractions are collected and stored in –70 °C. Small aliquots of crude extract, flow through, wash and each eluate were collected for analysis by SDS-PAGE (section 2.2.3.1).

2.2.2 Expression and Purification of recombinant *Drosophila* ADAR dsRNA binding domains in *E. coli*

2.2.2.1 pET28a expression vector

PCR products encoding the dsRNA binding domain were digested with *Nhe*I and *Sac*I and ligated into the corresponding restriction enzyme sites of the digested pET28a vector. The ligated DNA was transformed in *E. coli* DH5α cells and spread on an LB agar plate containing 50 μg/ml kanamycin. Plates were incubated overnight at 37 °C. Single colonies are selected and grown overnight in 2 ml LB medium containing 50 μg/ml kanamycin. Plasmid DNA was purified and an aliquot digested with *Eco*RI and *Kpn*I to check for inserts. Positive clones were sequenced to confirm that the construct was

correct and that the correct open reading frame fusing ADAR to the carboxy-terminal 6 X HIS epitope tag is present. Positive clones were transformed using standard methods into *E. coli* BL21 (DE3) prior to protein expression.

2.2.2.2 Expression and harvest of the dADAR dsRNA binding domains in *E. coli*

The pET28a plasmid containing the insert was transformed into electro competent BL21 (DE3) *E.coli* cells (section 2.1.3), plated on LB agar containing 50 μg/ml kanamycin and incubated overnight at 37 °C. A single bacterial colony was picked and 10 ml of M9 medium (200 ml 5x M9 salts, 2 ml 1 M $MgSO_4$, 100 ml 1 M $CaCl_2$, 10 ml 40% glucose, up to 1 litre with dH_2O) containing 50 μg/ml kanamycin was inoculated and grown overnight at 37 °C. 500 ml of fresh M9 medium containing 50 μg/ml kanamycin was inoculated with the 10 ml starter culture and grown at 37 °C to an OD_{600} of 0.6. At this point expression was induced with 500 μl of 1 M IPTG and the cultures grown for a further 4 hours. The cells were harvested by centrifugation. At this point the cells can be broken immediately or stored in –70 °C.

2.2.2.3 Extraction and purification of the dADAR dsRNA binding domains

The cells pellets were resuspended in 1 mL/g of sonication buffer with fresh protease inhibitors (50 mM KH_2PO_4, 500 mM NaCl, 20 mM imidazole, 0.5 mM PMSF, 0.7 mg/ml pepstatin, 0.4 mg/ml leupeptin) and sonicated in three 10 second bursts. Between sonication bursts the cells were incubated on ice for 20 seconds. The samples were centrifuged and the supernatant transferred to a fresh tube. At this point it is recommended to continue with purification rather than freezing the lysate, to avoid proteolysis. The dsRNA binding domains were purified over a Ni^{2+}-NTA column as described in section 2.2.1.8.

2.2.3 Analysis of recombinant proteins

2.2.3.1 Resolution of proteins on denaturing gels

Protein samples were mixed with 4x Laemmli Buffer (40% glycerol, 0.02% bromophenol blue, 4% β-mercaptoethanol, 8% SDS, 250 mM Tris-HCl, pH 6.8) and heated for 5 min at 95 °C before loading. Denaturing acrylamide minigels (10-12% acrylamide, 0.375 M Tris-HCl, pH 8.8, 0.1% SDS, 0.1% Ammonium persulphate, 0.2% TEMED) with stacking gels (4% acrylamide, 0.125 M Tris-HCl, pH 6.8, 0.1% SDS, 0.1% Ammonium persulphate, 0.2% TEMED) were poured using 30% acrylamide (29:1 acrylamide:bis-acrylamide, Severn Biotech). Gels were resolved in electrophoresis tanks (Pharmacia) in Tris-Glycine running buffer (25 mM Tris, 200 mM Glycine, pH 8.3, 0.1% SDS w/v) at 110 V for 2 hours. Broad range protein markers (Bio-Rad) were loaded to aid with analysis

2.2.3.2 Staining denaturing protein gels

Resolved protein gels had stacking gels removed and were then submerged in Coomassie stain (50% methanol, 10% acetic acid, 0.05% Coomassie blue R250, 40% dH_2O) for 1 hour with gentle agitation. The stain was discarded and gels were incubated for 1 hour in destain (20% isopropanol and 7.5% acetic acid) with gentle agitation.

2.2.3.3 Western Blotting

After resolution on denaturing gels protein samples were transferred to nitrocellulose for further analysis by semi-dry Western Blotting. Briefly, gels had the stacking gel removed and transfer apparatus (Bio-Rad) was assembled in accordance with the manufacturers' instructions with the gel and nitrocellulose membrane sandwiched on each side by 3 pieces of 3MM paper equilibrated in transfer buffer (50 mM Tris, 40 mM glycine, 0.05% SDS, 20% methanol). The apparatus was run at 120mA for 1 hour. Successful

transfer of protein was assessed by Ponceau S staining of the membrane – membranes were incubated with Ponceau S stain for 2 min. After two rinses with dH_2O membranes were visualised.

Specific proteins were detected on western blot membranes using antibody detection. Membranes were blocked for 1 hour with agitation in 1x PBS (phosphate buffered saline) with 5% milk protein (Marvel) or alternatively with a 1:10 dilution of Western Blocking Reagent (Roche) in PBS.

Primary antibodies were diluted in 5% milk, 1x PBS and incubated with membranes overnight at 4 °C or in a 1:20 dilution of Western Blocking Reagent (Roche) in PBS (antibody dilutions are shown in table 2.3). Membranes were washed 3 times in 1x PBS plus 0.05% Tween and incubated with secondary antibodies for at least 1 hour. After a further 3 washes membranes were detected with ECL (Amersham) and signals were exposed on Kodak Biomax film.

Table 2.3 Primary antibody dilutions for Western Blotting

Antibody	Species	Source	Dilution Factor
Anti-flag M2	Mouse	Sigma	1:3000
Anti-tetra HIS	Mouse	Qiagen	1:1000

2.2.3.4 Bradford Assay

Standard 100 µl BSA solutions were prepared of 1 µg/ml, 2 µg/ml, 5 µg/ml, 10 µg/ml, 15 µg/ml and 20 µg/ml from a stock solution of 0.5 mg/ml BSA. The protein sample to be measured was diluted in 100 µl dH_2O as appropriate, making sure a note of the dilution was taken. One part Protein

assay solution (Biorad) was mixed with four parts dH_2O and 900 μl of this mix was added to each of the standards and left at room temperature for 15 min. Each sample was measured on the spectrophotometer at 595 nm and a note taken of the reading. A standard graph is drawn using the values for the BSA standards and the concentration of the unknown sample is extrapolated accordingly.

2.3 Biochemical Assays

2.3.1 RNA methods

2.3.1.1 *In vitro* transcription of dsRNA

The dsRNA substrate used was ΔKP, which comprises exons 2 and 3 of the rat α-tropomyosin gene with no intron between. The sense RNA strand was synthesised using SP6 RNA polymerse (Stratagene) following linearisation of ΔKP with BamHI to give a 296 nt transcript. The antisense RNA was synthesised using T7 RNA polymerase (Stratagene) from a template generated by PCR using the SP6 and T7 promoter sequences as primers (See section 2.1.6.2). The reaction mix was as follows: 1 μg DNA template, 1U T7 or SP6 RNA polymerase (Stratagene) 5x transcription buffer, 10 mM DTT, 2 mM nCGU, 0.1 mM nATP, 150 μg/ml BSA, 1U RNA guard (Pharmacia), and if required, 10 μCi [α-^{32}P] ATP (3000 Ci/mmol; Amersham). The reaction was brought up to 25 μl with dH_2O and incubated at 37 °C for 1-2 h.

The sense and antisense strands were mixed at a molar ratio of 1:2 (sense:antisense) in buffer Q (50 mM Tris-HCl, pH 7.9, 100 mM KCl, 5 mM EDTA, 10% glycerol) and then heated to 75 °C for 10 min. The RNAs were allowed to cool slowly to 30 °C.

2.3.1.2 *In vitro* transcription of dADAR *exon 7*

The dADAR *exon 7* sense strand was synthesised using SP6 polymerase (Stratagene) from a template generated by linearisation of pGEM®-T Easy with *Nco*I. The reaction mix was as follows: 1 μg DNA template, 1U SP6 RNA polymerase (Stratagene) 5x transcription buffer, 10 mM DTT, 2 mM nCGUA, 150 μg/ml BSA, and 1U RNA guard (Pharmacia). The RNA was trace labelled with [α-^{32}P] ATP (3000 Ci/mmol; Amersham). The reaction was brought up to 25 μl with dH_2O and incubated at 37 °C for 1-2 h.

2.3.1.3 5′ End-labelling of primers

The *Adar exon 7* RT primer CAG TGT CGT TTT CGT GAG was 5′ end-labelled prior to poison primer extension. 5 pmol of primer was incubated with 1x PNK buffer, 10U polynuleotide kinase, 30 μCi [γ–^{32}P] dATP at 37 °C for 40 min. The reaction was stopped with 15 mM EDTA and 0.1% SDS and the RNA purified using a Sephadex G-25 spin column.

2.3.2 Editing Assays

2.3.2.1 Non-specific editing assay

Using 200 femto moles per reaction, the amount of radiolabelled dsRNA to set up the required number of reactions, including one reaction with no ADAR protein, was calculated. The dsRNA was ethanol precipitated, centrifuged, and the pellet washed with cold 70% ethanol. After drying in a spee vac, the pellet was then resuspended in 12.5 μl per reaction of 5 mM EDTA, 0.75 μl 5 mg/ml tRNA per reaction and 0.5 μl RNAguard (Pharmacia) per reaction.

Protein samples containing 1-12.5 μl of ADAR-containing protein fraction were set up on ice; buffer Q/200 was used to make up the remaining volume to 12.5 μl. 12.5 μl of the resuspended dsRNA substrate was added to each

EDTA, 0.75 µl 5 mg/ml tRNA per reaction and 0.5 µl RNAguard (Pharmacia) per reaction.

Protein samples containing 1-12.5 µl of ADAR-containing protein fraction were set up on ice; buffer Q/200 was used to make up the remaining volume to 12.5 µl. 12.5 µl of the resuspended dsRNA substrate was added to each protein sample and incubated for 1 hour at 37 °C. 8.3 µl of 7.5 M ammonium acetate and 120 µl of ice cold ethanol were added and the samples centrifuged at 12,000 x *g* for 30 min at 0 °C. The pellets were washed with 70% ethanol and centrifuged for a further 15 min at 0 °C before being vacuum dried. The dried pellets were resuspended in 10 µl of P1 buffer [30 mM potassium acetate pH 5.3, 10 mM zinc sulphate] containing 1.5 µg of 20 mg/ml nuclease P1 and incubated for 1 hour at 50 °C.

A 20 cm X 20 cm TLC plate was prepared by marking an origin line lightly with a pencil 1.5 cm from the bottom of the plate. Vertical lanes were created (15 lanes 1.3 cm wide or thirteen lanes 1.5 cm wide) by using the edge of a spatula to remove the cellulose layer between the lanes and unlabelled inosine was spotted at the origin of each lane and allowed to dry.

5 µl of the P1 nuclease digest was loaded at the origin of each lane on the TLC plate, allowed to dry before the second 5 µl was loaded. IMP was also spotted as a standard for inosine in each lane. The nucleotide mixture was resolved using 100 ml of chromatography solvent (saturated ammonium sulphate, 100 mM sodium acetate pH 6, and isopropanol (79:19:2)).

The plate was air-dried and exposed overnight to a PhosphorImager screen. 5′ α–^{32}P inosine migrates faster in this system than 5′ α–^{32}P adenosine. The radioactivity in the inosine and adenosine spots in each lane can be quantitated using ImageQuant™ software (Amersham Biosciences).

2.3.2.2 Poisoned primer extension

Based on using 10 femto moles per reaction, the amount of radiolabelled dsRNA to set up the required number of reations was calculated. As in section 2.3.1.1, the radiolabelled RNA substrate was ethanol precipitated, washed and dried and the pellet was resuspended in 12.5 µl of reaction mix (5 mM EDTA, 0.125 µl of poly(A) RNA, 0.5 µl RNAguard (Pharmacia), 20 mM DTT, 5 µg BSA).

As before, protein samples containing 1-12.5 µl of ADAR-containing protein fraction were set up on ice and buffer Q/200 was used to make up the remaining volume to 12.5 µl. 12.5 µl of the resuspended RNA substrate was added to each protein sample and incubated for 1 hour at 37 °C. 25 µl of proteinase K mixture consisting of 1 µl of 10 mg/ml proteinase K and 25 µl of 2x PK Buffer (20 mM Tris-HCl, pH 7.6, 20 mM NaCl, 2 mM $MgCl_2$, 1% SDS) was added to each sample and incubated for 30 min at 50 °C. Each sample was made up to 100 µl with dH_2O and the RNA phenol/chloroform extracted, ethanol precipitated and the pellet air-dried.

2 pico moles of the reverse transcription primer were $\gamma -^{32}P$ and resuspended in 100 µl dH_2O. The RNA pellet from the editing reaction was resuspended in annealing mix (0.5 µl of the end-labelled primer, 2 µl of 5x RT buffer, 0.5 µl of 0.5 µM ddTTP, 0.5 µl each of 2 mM dCTP, dGTP and dATP, made up to 8.3 µl with dH_2O) and incubated overnight at 55 °C to allow efficient annealing of the primer. The following day 1.3 µl of MMLV reverse transcriptase (Stratagene) was added to each sample and incubated for 1 hour at 42 °C.

Each sample was then ethanol precipitated and the dried pellets resuspended in 3 µl of formamide loading buffer boiled for 5 min at 95 °C and placed on ice. The samples were resolved on a 15% TBE/Urea gel, 0.5 mm spacer and

the gel dried before exposure to a PhosphorImager screen. The radioactivity in each lane was quantitated using ImageQuant™ (Amersham Biosciences).

2.3.3 Filter Binding assays

Increasing concentrations of pure recombinant ADAR dsRBMs were mixed with dsRNA containing 200 fmoles of ^{32}P-labelled adenosine in 50 µl of buffer C (50 mM Tris-HCl, pH 7.9, 150 mM KCl, 5 mM EDTA, 20% glycerol, 1 mM DTT, 0.5 mM PMSF, 0.7 µg/ml pepstatin, 0.4 µg/ml leupeptin, 0.15 mg/ml tRNA, 0.2 mg/ml BSA) for 10 min at 37 °C. The samples were then immediately filtered on a 0.45 µm nitrocellulose membrane filter (Schleicher & Schuell) that had been pre-soaked in the same buffer. The filters were then washed twice with 500 µl of buffer C, dried and the radioactivity was measured in a scintillation counter in the presence of scintillation liquid (Scintran).

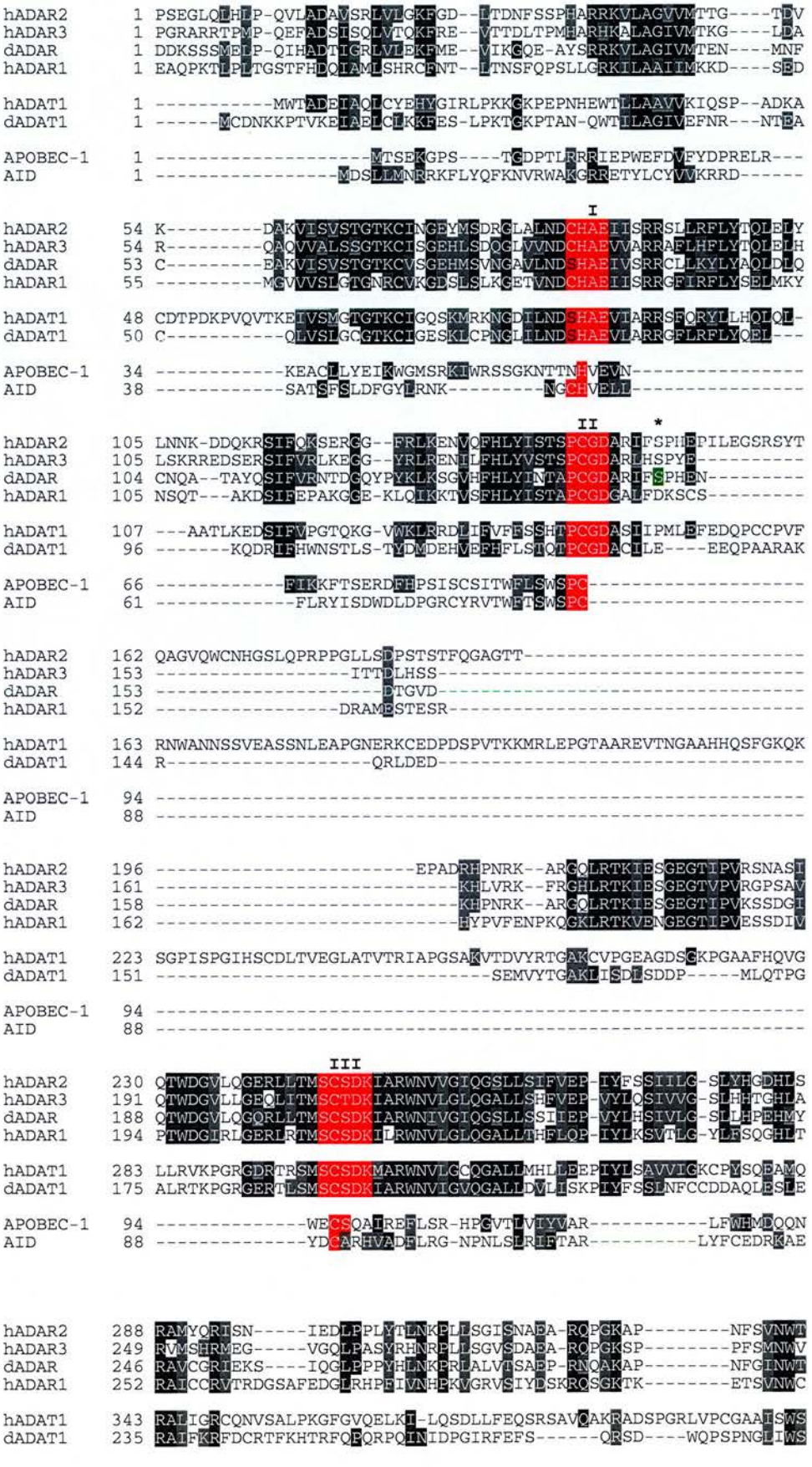
3

THE *DROSOPHILA* ADAR DEAMINASE DOMAIN

3.1 INTRODUCTION

The ADAR family of enzymes are a member of the larger family of cytidine deaminases (CDA). A C-terminal catalytic domain is found in all ADARs and contains conserved residues, $\underline{H}A\underline{E}(x)_{41\text{-}58}P\underline{C}G(x)_{44\text{-}154}S\underline{C}SDK$, analogous to sequences in the active sites of the cytidine deaminase family of enzymes. The cysteine and histidine residues are the ligands to the active site zinc, while the glutamate residue is necessary for proton transfer function. These motifs are also found in the ADAT (adenosine deaminase acting on tRNA) family of enzymes responsible for converting adenosine to inosine in tRNA. This homology between the ADARs and ADATs has prompted suggestions that the ADARs have evolved from the ADATs by the addition of dsRNA binding motifs.

Studies have shown that the deaminase domain of the ADAR enzymes is responsible for substrate specificity. I have cloned and purified the *Drosophila* deaminase domain and assayed its editing activity *in vivo* and *in vitro* on the dADAR self edit site, and on a dsRNA. The results suggest that the dADAR deaminase domain can edit dsRNA by itself. The editing pattern on the dsRNA is a subset of sites that are edited with the full length wild-type ADAR enzyme suggesting that site specificity is indeed conferred by the deaminase domain.


I
II
*
III

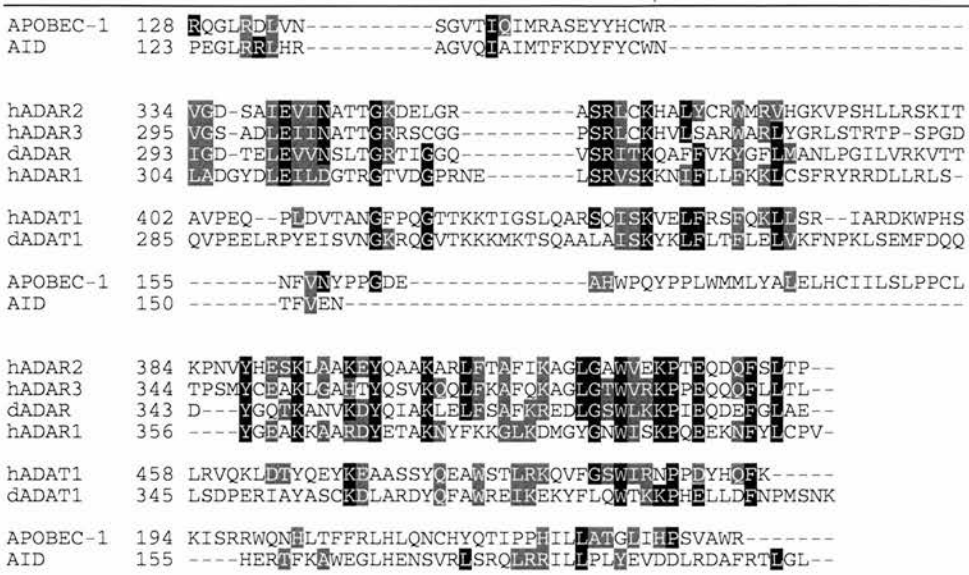


Figure 3.1 Sequence alignment of the *Drosophila* ADAR deaminase domain with the catalytic domains of other CDA family members.

The *Drosophila* ADAR deaminase domain is shown aligned by the Clustal W multiple sequence alignment program (http://www.ebi.ac.uk/clustalw/) with the catalytic domains of hADAR1, hADAR2, hADAR3, hADAT1, dADAT1, APOBEC-1 and AID. Conserved residues are highlighted in black with similar residues highlighted in gray. The three motifs, I, II and III, involved in the chelation of zinc and proton transfer are shown in red with the editing site of *Drosophila* ADAR indicated by an * and highlighted in green. Gaps are shown by dashed lines.

3.2 Alignment of ADAR catalytic domains

Amino acid sequence analysis of the *Drosophila* deaminase domain alongside the catalytic domains of other ADAR proteins has highlighted the presence of the three conserved CDA catalytic motifs and shown that dADAR is more similar to hADAR2 and hADAR3 than hADAR1. dADAR also shares good homology with both the human and *Drosophila* ADAT1 proteins. Figure 3.1 presents an alignment of the deaminase domains of a number of ADAR family and related proteins.

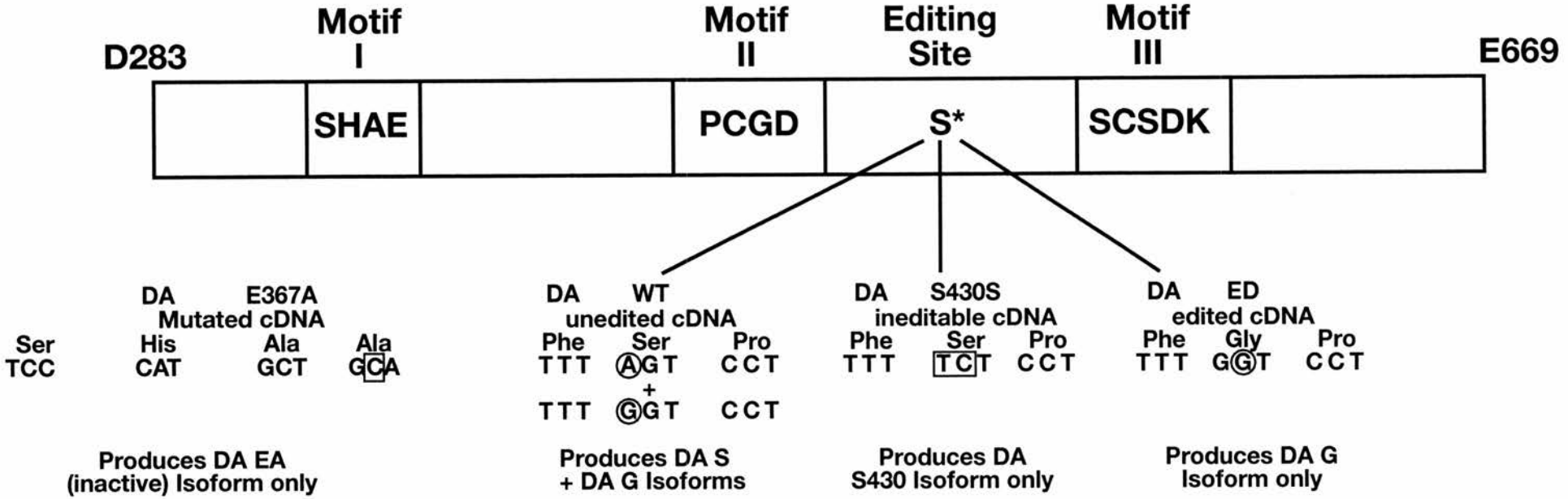
The ADATs are able to convert adenosine to inosine in tRNA with no identifiable RNA binding motif present. I decided to investigate whether the catalytic domain of dADAR was also capable of editing dsRNA with no identifiable RNA binding motifs present.

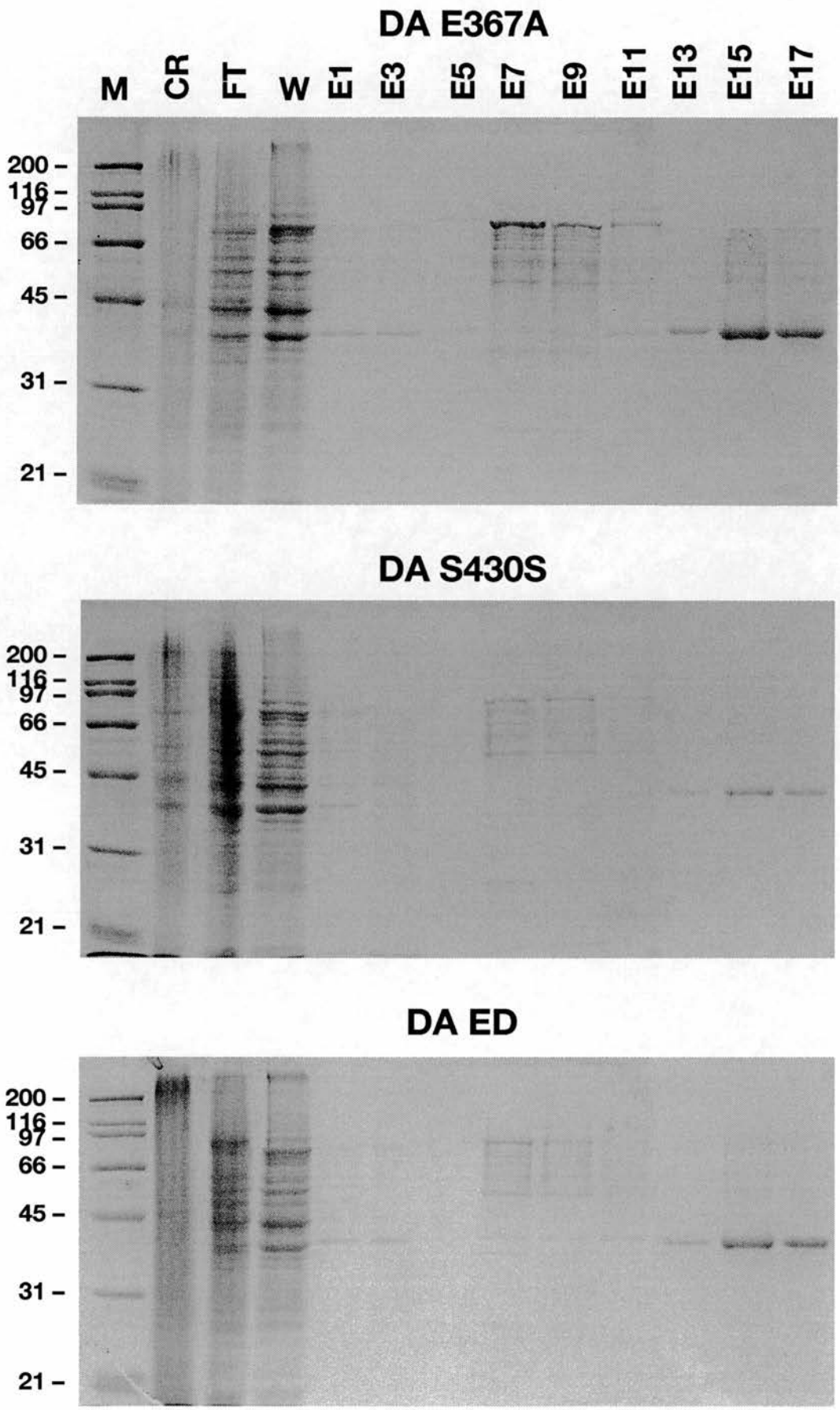
3.3 Expression and purification of dADAR deaminase domains

A number of constructs were designed based on previous studies of the dADAR full length enzyme. Wild-type dADAR is capable of editing its own transcript when expressed in *Pichia pastoris* resulting in a mix of edited and unedited isoforms. To allow the expression of a single isoform rather than a mixed population the edited serine codon AGT can be mutated to another codon for serine, the ineditable form TCT or alternatively, the AGT serine codon can be mutated to glycine GGT allowing the expression of the edited

Figure 3.2A Schematic of the *Drosophila* ADAR Deaminase Domain Constructs were designed to begin translation at aspartate residue 283 of the full length dADAR protein and a C-terminal tail containing the 6 X histidine tag was added. The point mutations required to produce the various isoforms are illustrated.

Drosophila ADAR Deaminase Domain (DA)




DA E367A
M
CR
FT
W
E1
E3
E5
E7
E9
E11
E13
E15
E17
200
116
97
66
45
31
21
DA S430S
200
116
97
66
45
31
21
DA ED
200
116
97
66
45
31
21

isoform alone. A point mutation in motif I changing the glutamate residue to alanine has been shown to result in an inactive isoform (Lai et al., 1995; Liu and Samuel, 1996). As all these modifications occur within the catalytic region I was therefore able to amplify and clone four different isoforms of the dADAR deaminase domain; wild-type, ineditable (S430S), edited (S430G)and inactive (E367A) (Figure 3.2a); into the pPICZa vector in frame with the 6x HIS tag. Once transformed into *Pichia* colonies the expression of these 6x HIS tagged proteins was efficient giving approximately 8 mg protein per 1 litre of culture after purification over Ni^{2+}-NTA agarose. SDS-PAGE and western blot analysis using an anti-tetra HIS antibody detected a 45 kDa band corresponding to the predicted molecular weight of the deaminase domain constructs (Figure 3.2b).

3.4 THE *DROSOPHILA* ADAR WILD-TYPE DEAMINASE DOMAIN DISPLAYS NO ACTIVITY *IN VIVO*

As discussed before, the full length dADAR enzyme is able to edit its own transcript at the exon 7 site when expressed in the yeast *Pichia pastoris*. Because editing at the *Adar exon 7* site does not require any intronic sequences to form a permissible structure the expression of a recombinant dADAR protein from a cDNA clone results in active protein and a high affinity site being produced within the yeast cell. As there is no endogenous

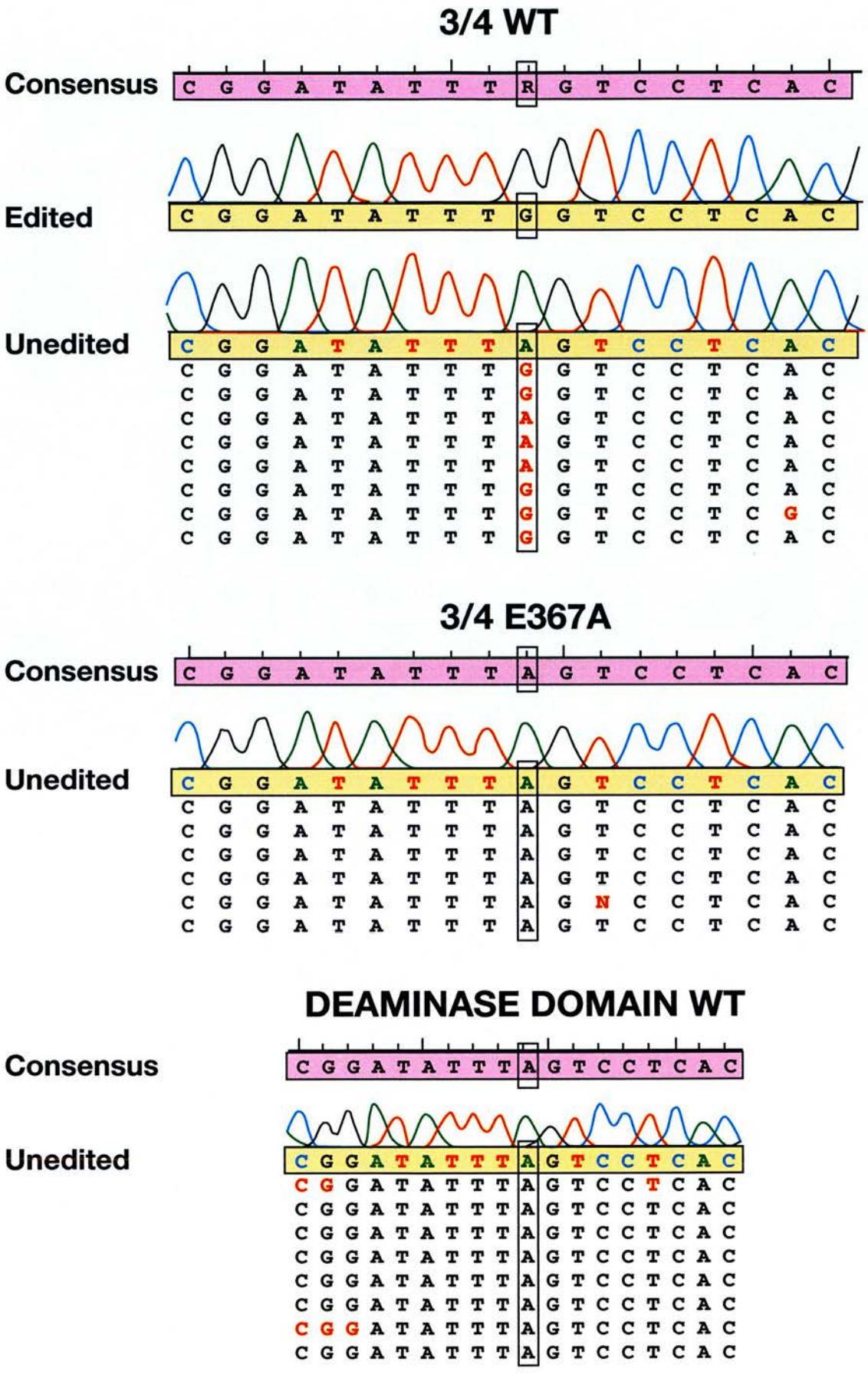
Figure 3.2B Purification of recombinant dADAR deaminase domain (DA) proteins. His-tagged DA proteins were purified from *Pichia pastoris* over Ni^{2+}NTA agarose. Samples were taken from the crude (CR), flow through (FT), wash (W) and elutions (E1-E17) and the level of purification was determined by SDS-PAGE and Coomassie staining. Protein markers (kDa) are indicated on the left.

ADAR activity in *Pichia* any A-to-I editing of mRNA must be due to the recombinant ADAR proteins expressed.

To assess the activity of the deaminase domain, RT-PCR and sequence analysis was performed on total RNA isolated from *Pichia pastoris* expressing either full length dADAR 3/4 WT (full length dADAR minus exon 3a), full length dADAR 3/4 E367A or the wild-type deaminase domain. As demonstrated by Palladino *et al* the 3/4 wild-type enzyme showed high levels of self editing (Palladino et al., 2000a) whilst the inactive E367A isoform displayed no activity (Figure 3.3). Unfortunately, the wild-type deaminase domain also showed no activity suggesting that the dsRNA binding domains may be required for activity (Figure 3.3).

Western blot analysis of the recombinant wild-type deaminase domain using an anti-tetra HIS antibody showed two bands of approximately 45 kDa and 96 kDa (Figure 3.4). To further confirm the purification of the deaminase domain the upper and lower bands from elution fraction 7 were separated by SDS-PAGE and the bands analysed using peptide mass fingerprinting. As a control a sample of active full length dADAR 3/4 S430S was also analysed. The two bands from the deaminase domain purification produced good spectra and the peptides detected post tryptic digest suggested they were the same protein however, theoretical digest on the *Drosophila* ADAR sequence did not match. Database searches with the peptide sequences did not produce any hits suggesting that rather than purifying the dADAR

Fig 3.3 Editing at the dADAR exon 7 self editing site RNA was isolated from individual *Pichia* clones expressing dADAR 3/4 wild-type (3/4 WT), dADAR 3/4 E367A (3/4 E367A) and wild-type deaminase domain were sequenced and multiple sequence chromatograms show the level of editing at the *Adar exon 7* self editing site for each enzyme. 3/4 WT shows 62% editing at exon 7 (n=13). For 3/4 E367A (n=7) and for deaminase domain WT (n=25) no editing activity was detected.


3/4 WT
Consensus
C G G A T A T T T R G T C C T C A C
Edited
C G G A T A T T T G G T C C T C A C
Unedited
C G G A T A T T T A G T C C T C A C
C G G A T A T T T G G T C C T C A C
C G G A T A T T T G G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T G G T C C T C A C
C G G A T A T T T G G T C C T C G C
C G G A T A T T T G G T C C T C A C
3/4 E367A
Consensus
C G G A T A T T T A G T C C T C A C
Unedited
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G N C C T C A C
C G G A T A T T T A G T C C T C A C
DEAMINASE DOMAIN WT
Consensus
C G G A T A T T T A G T C C T C A C
Unedited
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C
C G G A T A T T T A G T C C T C A C

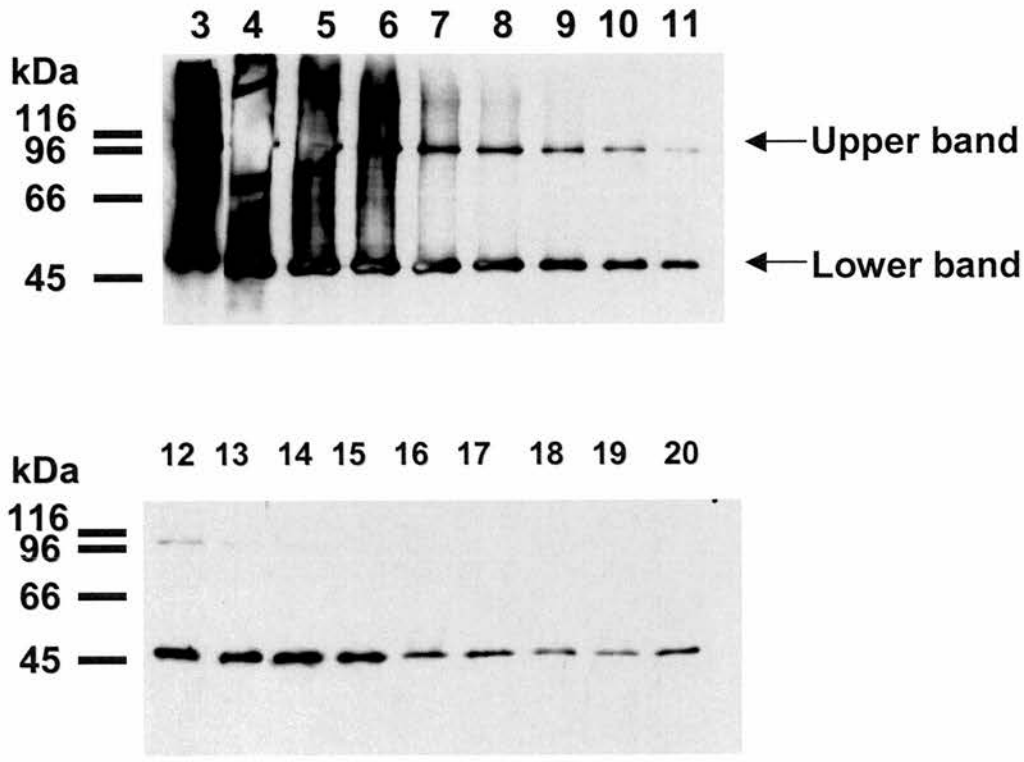

3 4 5 6 7 8 9 10 11
kDa
116
96
66
45
Upper band
Lower band
12 13 14 15 16 17 18 19 20
kDa
116
96
66
45

Figure 3.4 Purification of wild-type dADAR deaminase domain The elution fractions from the purification of wild-type deaminase domain over Ni^{2+} NTA agarose was analysed by Western blot using an anti-tetra HIS antibody (Qiagen). Two bands were apparent, a 45 kDa band matching the predicted size of the recombinant dADAR deaminase domain protein and an unknown band of approximately 96 kDa. The upper and lower bands from elution fraction 7 were analysed by peptide mass fingerprinting. Protein size markers (kDa) are indicated on the left.

deaminase domain I was instead purifying a *Pichia* contaminant of a similar size that cross reacted with the anti-tetra His antibody. It should be noted that the control sample of dADAR 3/4 S430S did map to the full length protein.

The peptide fingerprinting results explain the lack of editing seen in the RT-PCR sequence analysis with the deaminase sample. One explanation is that the deaminase domain construct was too short and the protein that was generated was unstable and was degraded. If this occurred then the contaminant *Pichia* protein would bind to and be purified from the Ni^{2+}-NTA column.

3.5 AN ALTERNATIVE DEAMINASE DOMAIN CONSTRUCT

To further confirm whether it is possible to produce recombinant dADAR deaminase domain, an alternative construct was cloned into pPICZ A and transformed in *Pichia*. This construct, termed DM WT, instead begins at P270 and has an N-terminal FLAG tag and a C-terminal histidine tag immediately downstream of the last glutamate residue (Gallo et al., 2003). SDS-PAGE and Western blot analysis using an anti-tetra HIS antibody detected a 45 kDa band corresponding to the predicted molecular weight of the DM WT

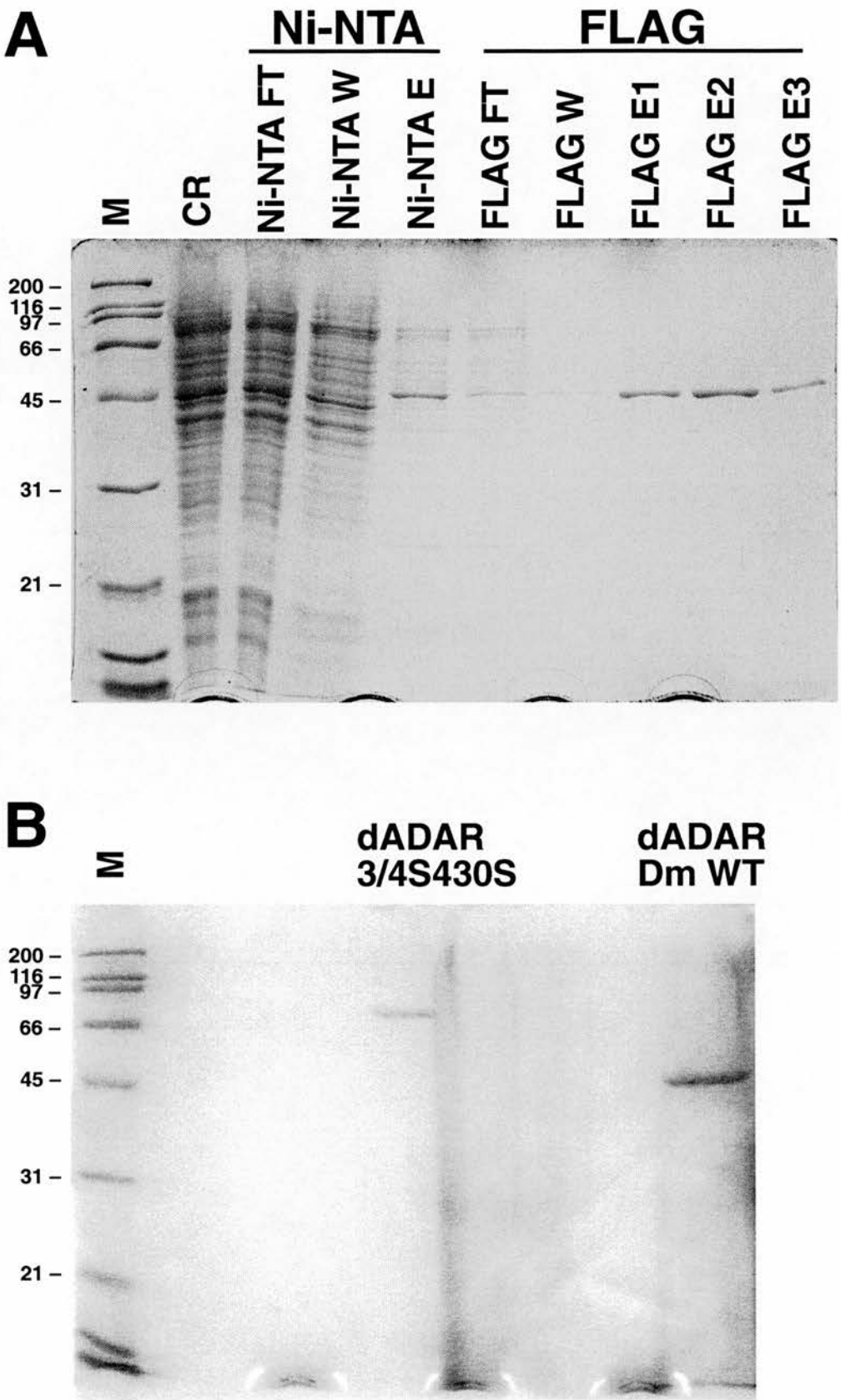

A
Ni-NTA
FLAG
M
CR
Ni-NTA FT
Ni-NTA W
Ni-NTA E
FLAG FT
FLAG W
FLAG E1
FLAG E2
FLAG E3
200 –
116 –
97 –
66 –
45 –
31 –
21 –
B
M
dADAR 3/4S430S
dADAR Dm WT
200 –
116 –
97 –
66 –
45 –
31 –
21 –

Figure 3.5A Purification of recombinant wild-type *Drosophila* ADAR deaminase domain (DM WT) protein. N-terminal FLAG-tagged and C-terminal His-tagged DM proteins were purified from *Pichia pastoris* over Ni^{2+}NTA and FLAG agarose. Samples were taken from the crude (CR), Ni^{2+}NTA flow through (Ni-NTA FT), Ni^{2+}NTA wash (Ni-NTA W), Ni^{2+}NTA eluate (Ni^{2+}-NTA E), FLAG flow through (FLAG FT), FLAG wash (FLAG W) and FLAG elutions (FLAG E1, FLAG E2, FLAG E3) and the level of purification was determined by SDS-PAGE and Coomassie staining. Protein markers (kDa) are indicated on the left.

Figure 3.5B Peptide Mass fingerprinting of dADAR 3/4 S430S and DM WT proteins Samples from dADAR 3/4 S430S and DM WT purifications were separated by SDS-PAGE and the bands excised and analysed by peptide mass fingerprinting.

construct (Figure 3.5A). Once again a sample was run on SDS-PAGE and the band excised for peptide mass fingerprinting alongside the full length dADAR 3/4 S430S protein (Figure 3.5B). The peptides for the full length protein mapped the entire protein whilst the peptides from the DM WT digest mapped to the C-terminal portion of the protein (Figures 3.5C and 3.5D).

3.6 THE dADAR DM WT CAN EDIT *ADAR EXON 7 IN VIVO* IN *PICHIA PASTORIS*

Now that peptide mass fingerprinting had confirmed the expression and purification of DM WT I decided to assess its ability to deaminate its own transcript. As before, RT-PCR and sequence analysis was performed on total RNA isolated from *Pichia pastoris* expressing DM WT and of the 25 sequences examined, 3 showed editing at the exon 7 S/G editing site (Figure 3.6A).

On analysing the *exon 7* sequences it became apparent that a number of other editing events occur within this transcript. In order to clarify whether the

1 11 21 31 41 51 61 71

MLNSANNNSP QHPVSAPSDI NMNGYNRKLP QKRGYEMPKY SDPKKKMCKE RIPQPKNTVA MLNELRHGLI YKLESQTGPV

81 91 101 111 121 131 141 151

HAPLFTISVE VDGQKYLGQG RSKKVARIEA AATALRSFIQ FKDGAVLSPL KPAGNLDFTS DEHLENGIEN LSSSKMFEII

161 171 181 191 201 211 221 231

QTMLTEKLSN PTSLEQPTFC MSQNVSKSAI TVDGQKKVPD KGPVMLLYEL FNDVNFECIN IDGAQNNCRF KMTVTINEKK

241 251 261 271 281 291 301 311

FDGTGPSKKT AKNAAAKAAL ASLCNISYSP MVVPQKNVPL PIDDKSSSME LPQIHADTIG RLVLEKFMEV IKGQEAYSRR

321 331 341 351 361 371 381 391

KVLAGIVMTE NMNFCEAKVI SVSTGTKCVS GEHMSVNGAV LNDSHAEIVS RRCLLKYLYA QLDLQCNQAT AYQSIFVRNT

401 411 421 431 441 451 461 471

DGQYPYKLKS GVHFHLYINT APCGDARIFG PHENDTGVDK HPNRKARGQL RTKIESGEGT IPVKSSDGIQ TWDGVLQGQR

481 491 501 511 521 531 541 551

LLTMSCSDKI ARWNIVGIQG SLLSSIIEPV YLHSIVLGSL LHPEHMYRAV CGRIEKSIQG LPPPYHLNKP RLALVTSAEP

561 571 581 591 601 611 621 631

RNQAKAPNFG INWTIGDTEL EVVNSLTGRT IGGQVSRITK QAFFVKYGFL MANLPGILVR KVTTDYGQTK ANVKDYQIAK

641 651 661

LELFSAFKRE DLGSWLKKPI EQDEFGLAE

Figure 3.5c Post trypsin digest peptides from dADAR 3/4 S430S sample The peptides obtained from trypsin digest of the dADAR 3/4 S430S band mapped across the entire dADAR peptide sequence (Accession no. 7546826). The matched peptides cover **35%** (239/669 aa's) of the protein.

1 11 21 31 41 51 61 71
MLNSANNNSP QHPVSAPSDI NMNGYNRKLP QKRGYEMPKY SDPKKKMCKE RIPQPKNTVA MLNELRHGLI YKLESQTGPV
81 91 101 111 121 131 141 151
HAPLFTISVE VDGQKYLGQG RSKKVARIEA AATALRSFIQ FKDGAVLSPL KPAGNLDFTS DEHLENGIEN LSSSKMFEII
161 171 181 191 201 211 221 231
QTMLTEKLSN PTSLEQPTFC MSQNVSKSAI TVDGQKKVPD KGPVMLLYEL FNDVNFECIN IDGAQNNCRF KMTVTINEKK
241 251 261 271 281 291 301 311
FDGTGPSKKT AKNAAAKAAL ASLCNISYSP MVVPQKNVPL PIDDK**SSSME LPQIHADTIG** RLVLEKFMEV IKGQEAYSRR
321 331 341 351 361 371 381 391
K**VLAGIVMTE NMNFCEAK**VI SVSTGTKCVS GEHMSVNGAV LNDSHAEIVS RRCLLKYLYA QLDLQCNQAT AYQSIFVRNT
401 411 421 431 441 451 461 471
DGQYPYKLK**S GVHFHLYINT APCGDAR**IFG PHENDTGVDK HPNRKARGQL RTKIESGEGT IPVK**SSDGIQ TWDGVLQGQR**
481 491 501 511 521 531 541 551
LLTMSCSDKI ARWNIVGIQG SLLSSIIEPV YLHSIVLGSL LHPEHMYRAV CGRIEK**SIQG LPPPYHLNKP RLALVTSAEP**
561 571 581 591 601 611 621 631
RNQAK**APNFG INWTIGDTEL EVVNSLTGR**T IGGQVSRITK QAFFVK**YGFL MANLPGILVR** KVTTDYGQTK ANVKDYQIAK
641 651 661
LELFSAFKRE DLGSWLKKPI EQDEFGLAE

Figure 3.5d Post trypsin digest peptides from DM WT sample The peptides obtained from trypsin digest of the DM WT band mapped onto the carboxy portion of the dADAR peptide sequence (Accession no. 7546826). The matched peptides cover **21%** (147/669 aa's) of the protein.

DM WT

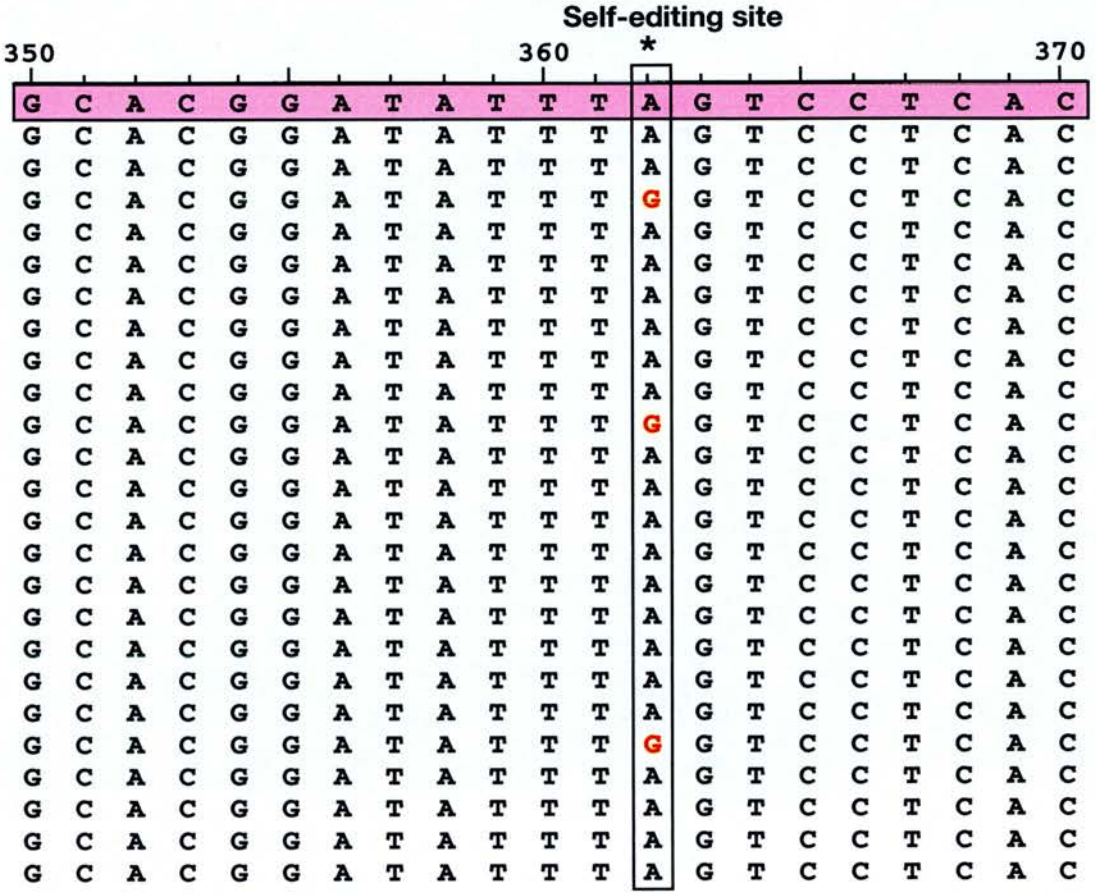
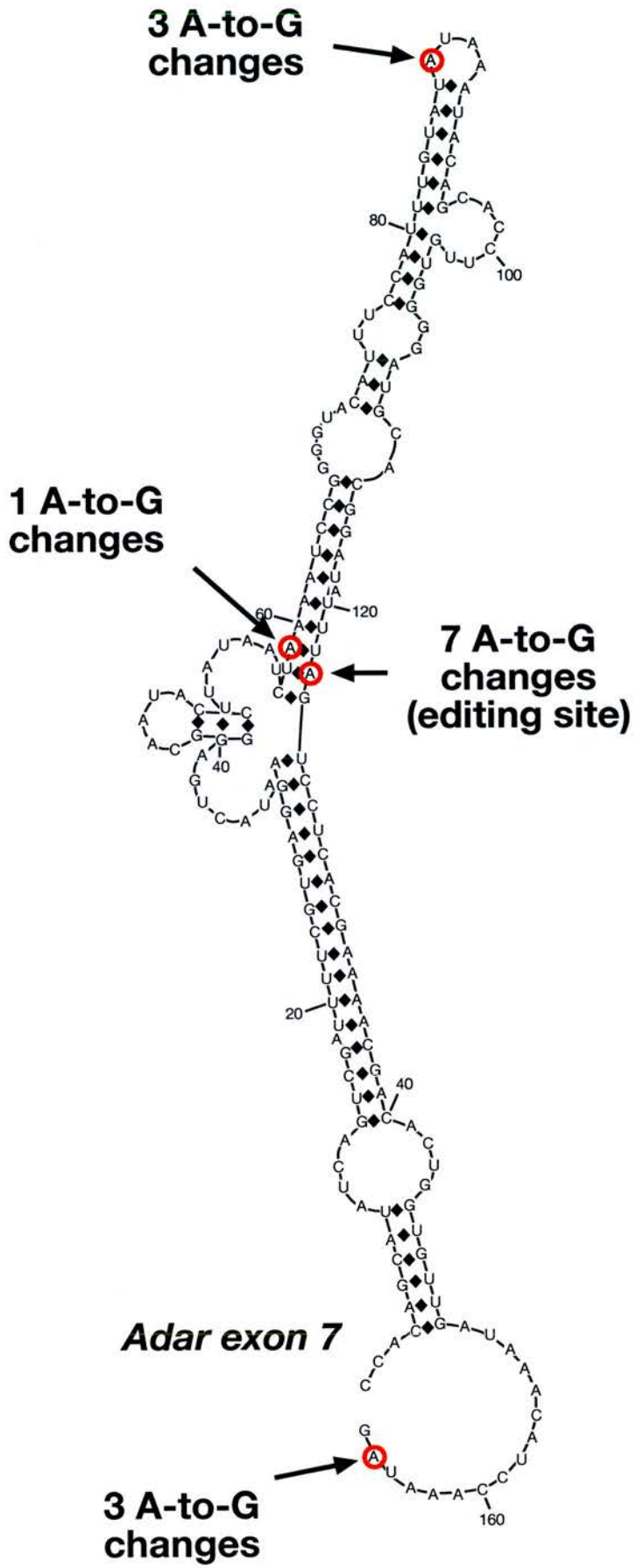
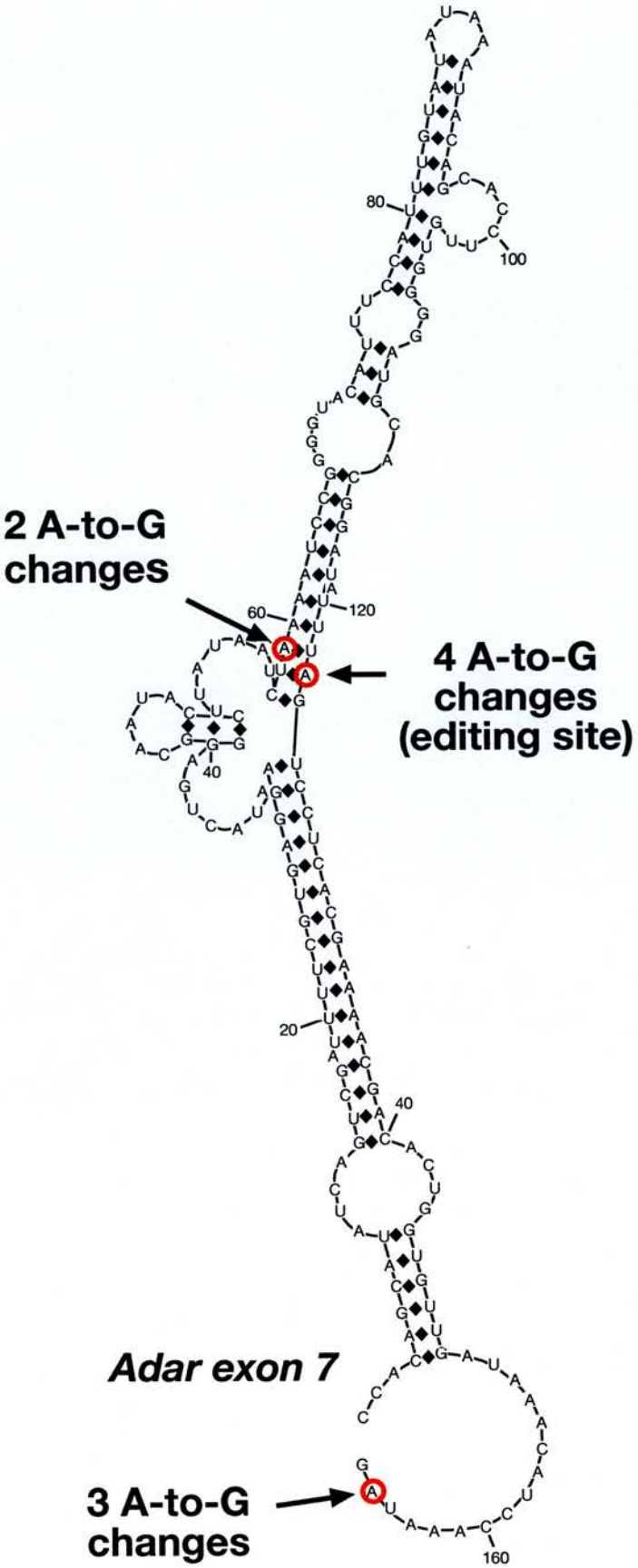

Self-editing site
350
360
*
370
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T G G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T G G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T G G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C
G C A C G G A T A T T T A G T C C T C A C

Figure 3.6A Editing at the dADAR exon 7 self editing site by DM WT RNA was isolated from individual *Pichia* clones expressing dADAR DM WT were sequenced and multiple sequence chromatograms show the level of editing at the *Adar exon 7* self editing site. DM WT shows 12% editing at exon 7 (n=25).

Figure 3.6Bi Mapping of dADAR 3/4 S430S editing on the exon 7 substrate. RNA structure prediction for all 187 bases of exon 7 of the *Adar* transcript using the *mfold* program (Zuker, 1989). 13 *Adar* transcripts isolated from individual *Pichia* clones were analysed and all A-to-G changes mapped. Full length dADAR 3/4 S430S can edit exon 7 of the *Adar* transcript at 4 different sites.

Figure 3.6Bii Mapping of dADAR DM WT editing in the exon 7 substrate. RNA structure prediction for all 187 bases of exon 7 of the *Adar* transcript using the *mfold* program (Zuker, 1989). 33 *Adar* transcripts isolated from individual *Pichia* clones were analysed and all A-to-G changes mapped. DM WT can edit exon 7 of the *Adar* transcript at 3 different sites.

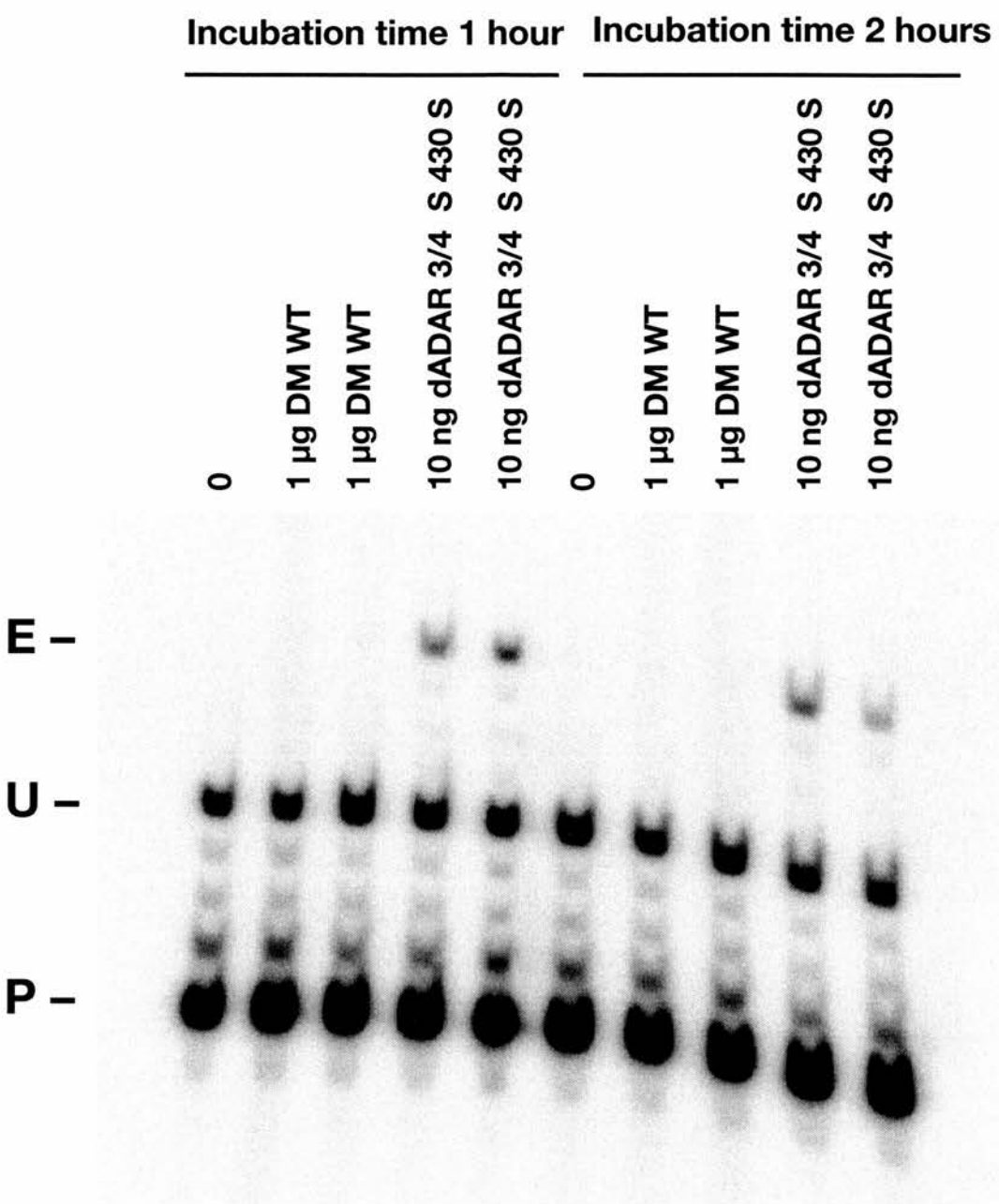

3 A-to-G changes
1 A-to-G changes
7 A-to-G changes (editing site)
Adar exon 7
3 A-to-G changes
20
40
60
80
100
120
160


2 A-to-G changes
4 A-to-G changes (editing site)
Adar exon 7
3 A-to-G changes
20
40
60
80
100
120
160

deaminase domain did indeed confer a level of site specificity I decided to map the editing events within *Adar exon 7* and compare the sites targeted by full length dADAR 3/4 WT and the DM WT. DM WT and 3/4 WT were expressed in *Pichia pastoris,* the total RNA extracted and RT-PCR performed. The resulting PCR products were TA cloned into pGEM®-T Easy vector and transformed into *E. coli.* Colonies were screened and plasmid preps made before sequencing using the SP6 and T7 promoter sequences within the pGEM®-T Easy multiple cloning site.

Four editing sites were identified, including the well-documented S/G site. With reference to the RNA fold illustrated in Figure 3.6Bi, of the 13 transcripts analysed full length dADAR targeted four adenosines within exon 7 including the S/G site at 54%, two sites situated within loops at either end of the structure (A87 and A165) at 23% and another adenosine (A59) immediately opposite the S/G site at 8%. A far greater number of transcripts were sequenced to map the editing of DM WT indicating it edits much less efficiently than full length dADAR. Of the 33 sequences analysed (Figure 3.6Bii) DM WT was shown to target three adenosines within exon 7; the S/G site at 11%, A165 at 3% and A59 immediately opposite the S/G site at 6%.

Figure 3.7 Poisoned primer extension assay comparing editing *in vitro* on *Adar exon 7* by DM WT and dADAR 3/4 S430S. DM WT and dADAR 3/4 S430S were incubated with 10 femto moles *Adar exon 7* transcript for 1 and 2 hours. The first lane contains no ADAR protein. Lanes 4 and 5 are duplicate lanes containing 125 pico moles 3/4 S430S that edits *Adar exon 7* transcript to 20% after one hour incubation rising to 25% after 2 hours. 22 n moles DM WT showed no editing activity on *Adar exon 7 in vitro.* Editing activity was quantified using a phosphoimager and ImageQuant™ software. Radiolabelled primer; P, the unedited; U and edited; E primer extensions products.



DM vs exon7 PE Assay

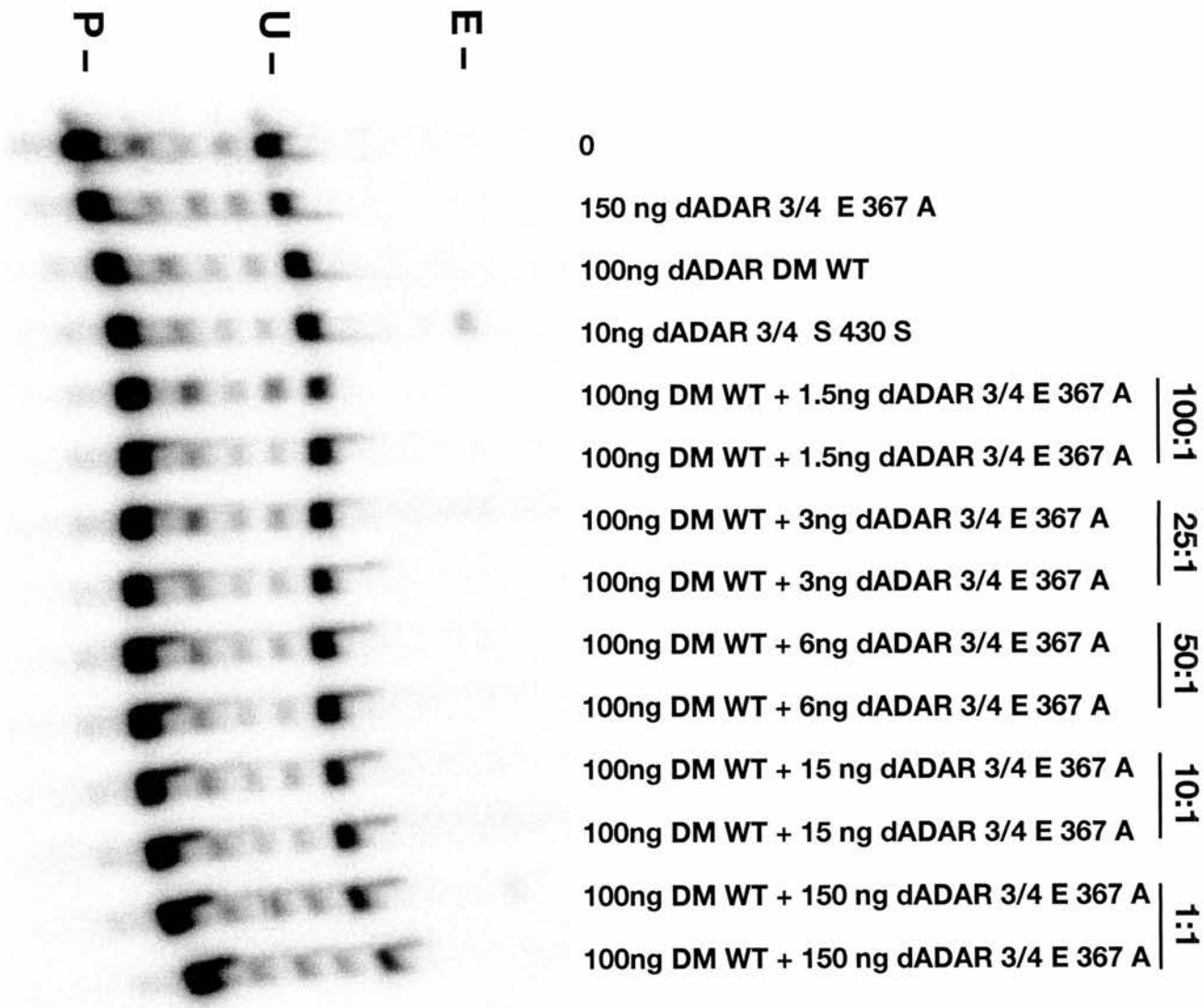
This overlap of specificity between full length ADAR and the deaminase domain indicates that the catalytic domain does indeed confer a level of specificity though without the dsRNA binding domain it edits far less efficiently than the full length protein.

3.7 THE dADAR DM WT CAN NOT EDIT *ADAR EXON 7 IN VITRO*

Now DM WT has shown it is capable of editing dsRNA I performed a poisoned primer extension assay to measure the DM WT editing activity on the S/G site in *exon 7* transcripts *in vitro*. Full length 3/4 S430S edited exon 7 at about 20% but no DM WT editing activity was detected (Figure 3.7).

Macbeth *et al.* recently reported the crystal structure of the hADAR2 deaminase domain. The construct analysed was similar to DM WT (See appendix 2) and able to edit dsRNA substrates *in vitro* but was less efficient than full length hADAR2 (Macbeth et al., 2005). In the case of dADAR full length 3/4 S430S will normally edit *exon 7 in vitro* to approximately 70% therefore in this case the control was working at a much lower efficiency than normally seen. If DM WT is able to edit this transcript *in vitro* levels may be below the threshold of the experiment.

Figure 3.8 Poisoned primer extension assay on *Adar exon 7* editing *in vitro* by DM WT and increasing amounts of dADAR 3/4 E367A 2.2 nano moles DM WT and increasing amounts (18.75 pico moles to 1875 pico moles) of dADAR 3/4 E367A were incubated with 10 femto moles *Adar exon 7* transcript. The first lane contains no ADAR protein. Lane 2 contains dADAR 3/4 E367A and lane 3 contains DM WT; neither demonstrated editing activity on *Adar exon 7*. Lane 4 shows editing of *Adar exon 7* transcript to 5% by dADAR 3/4 S430S. Lanes 5 to 14 contain constant amounts of DM WT with increasing amounts of dADAR 3/4 E367A added. No editing was seen. Editing activity was quantified using a phosphoimager and ImageQuant™ software. Radiolabelled primer; P, the unedited; U and edited; E primer extensions products.

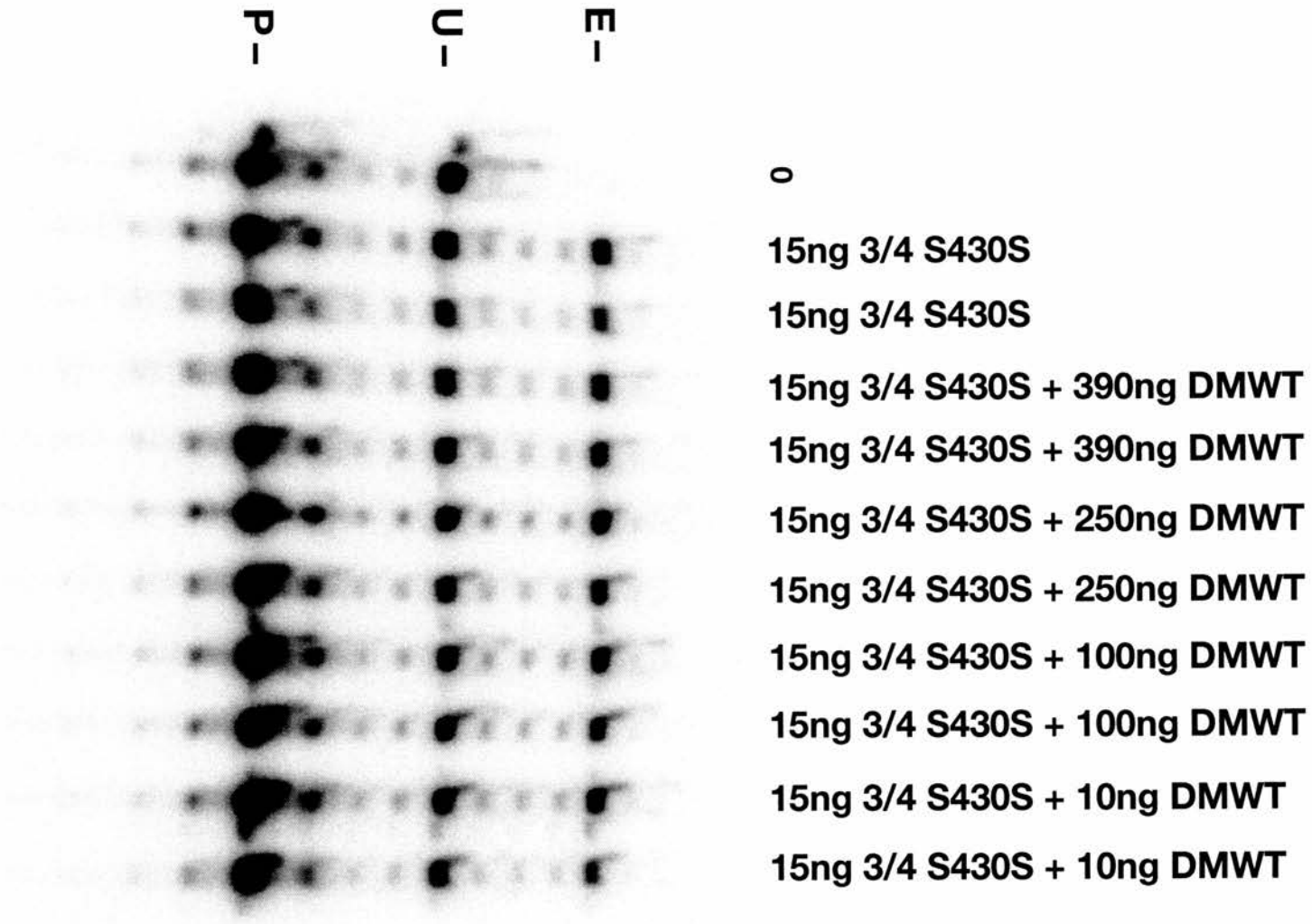

P –
U –
E –
0
150 ng dADAR 3/4 E 367 A
100ng dADAR DM WT
10ng dADAR 3/4 S 430 S
100ng DM WT + 1.5ng dADAR 3/4 E 367 A
100ng DM WT + 1.5ng dADAR 3/4 E 367 A
100:1
100ng DM WT + 3ng dADAR 3/4 E 367 A
100ng DM WT + 3ng dADAR 3/4 E 367 A
25:1
100ng DM WT + 6ng dADAR 3/4 E 367 A
100ng DM WT + 6ng dADAR 3/4 E 367 A
50:1
100ng DM WT + 15 ng dADAR 3/4 E 367 A
100ng DM WT + 15 ng dADAR 3/4 E 367 A
10:1
100ng DM WT + 150 ng dADAR 3/4 E 367 A
100ng DM WT + 150 ng dADAR 3/4 E 367 A
1:1

To test whether another dsRNA binding protein might facilitate the catalytic function of the deaminase domain, DM WT was incubated with *Adar exon 7.* and increasing amounts of full length 3/4 E367A. Previous studies have shown that the deaminase domain does not interact with or form dimers with full length dADAR protein (Gallo et al., 2003).

However, the conformational changes induced by the binding of a full length ADAR protein (Yi-Brunozzi et al., 2001) may facilitate editing by the deaminase domain. Therefore any editing that occurred would be the function of the DM WT protein as the E367A mutation in the full length dADAR protein renders it unable to deaminate dsRNA yet it retains its ability to bind to dsRNA. No editing by DM WT was seen (Figure 3.8) suggesting it did not interact with the 3/4 E367A protein.

If the activity of DM WT can not be supplemented by the presence of a full length dADAR protein then perhaps it may inhibit the activity of an active full length ADAR isoform. 3/4 S430S was incubated with *Adar exon 7* transcript and decreasing amounts of DM WT. The level of editing by 3/4 by 3/4 S430S alone was first determined to be 36%. The addition of DM WT did not alter this significantly and levels remained steady at approximately 40%.

Figure 3.9 Poisoned primer extension assay on *Adar exon 7* editing *in vitro* by dADAR 3/4 S430S and decreasing amounts of DM WT. 187.5 pico moles dADAR 3/4 S430S and decreasing amounts (8.67 nano moles to 222 pico moles) of DM WT were incubated with 10 femto moles *Adar exon* 7 transcript. The first lane contains no ADAR protein. Lanes 2 and 3 are duplicate lanes containing 3/4 S430S that edits *Adar exon 7* transcript to 36%. Lanes 4 to 11 contain constant amounts of 3/4 S430S with decreasing amounts of DM WT added. Editing of *Adar exon 7* transcript in these lanes remains constant at approximately 40%. Editing activity was quantified using a phosphoimager and ImageQuant™ software. Radiolabelled primer; P, the unedited; U and edited; E primer extensions products.


P–
U–
E–
0
15ng 3/4 S430S
15ng 3/4 S430S
15ng 3/4 S430S + 390ng DMWT
15ng 3/4 S430S + 390ng DMWT
15ng 3/4 S430S + 250ng DMWT
15ng 3/4 S430S + 250ng DMWT
15ng 3/4 S430S + 100ng DMWT
15ng 3/4 S430S + 100ng DMWT
15ng 3/4 S430S + 10ng DMWT
15ng 3/4 S430S + 10ng DMWT

The presence of full length dADAR proteins with which the possibility of an interaction via the deaminase domain may exist, have no effect on DM WT activity with the reciprocal being true in that DM WT has no effect on their enzymatic activity.

3.8 THE dADAR DM WT CAN EDIT NON-SPECIFIC dsRNA *IN VITRO*

If DM WT is unable to edit a specific dsRNA substrate *in vitro* perhaps a dsRNA duplex would provide a more accessible substrate. I assayed DM WT activity on the dsRNA substrate ΔKP, comprising exons 2 and 3 of the rat α-tropomyosin gene with no intron between, and analysed the results by thin layered chromatography. Under conditions where editing of between 28% and 48% were obtained for increasing amounts of 3/4 S430S, DM WT showed no editing activity. This method of analysis is not very sensitive therefore very low levels of editing would remain undetected.

I decided to analyse the ΔKP substrate by RT-PCR and sequencing each strand. The sense strand was reverse transcribed using the KP 3′ Sp6 oligo and the anti sense strand using the Sp6 promoter oligo. The sense strand was then amplified using KP 3′ Sp6 and KP 5′ Sp6, with the anti sense strand amplified by the Sp6 promoter oligo and KP 5′ T7 (see Table 2.1 for oligo sequences). The PCR products were TA cloned into the pGEM®-T Easy vector as before and positive clones sequenced.

Analysis of the sequences demonstrated that DM WT did indeed edit the ΔKP dsRNA substrate at a number of sites. As before the levels of editing were much lower than that seen for the full length 3/4 S430S protein however, again reiterating the earlier results on *Adar exon 7*, there was a definite specificity in the adenosines targeted with DM WT targeting fewer

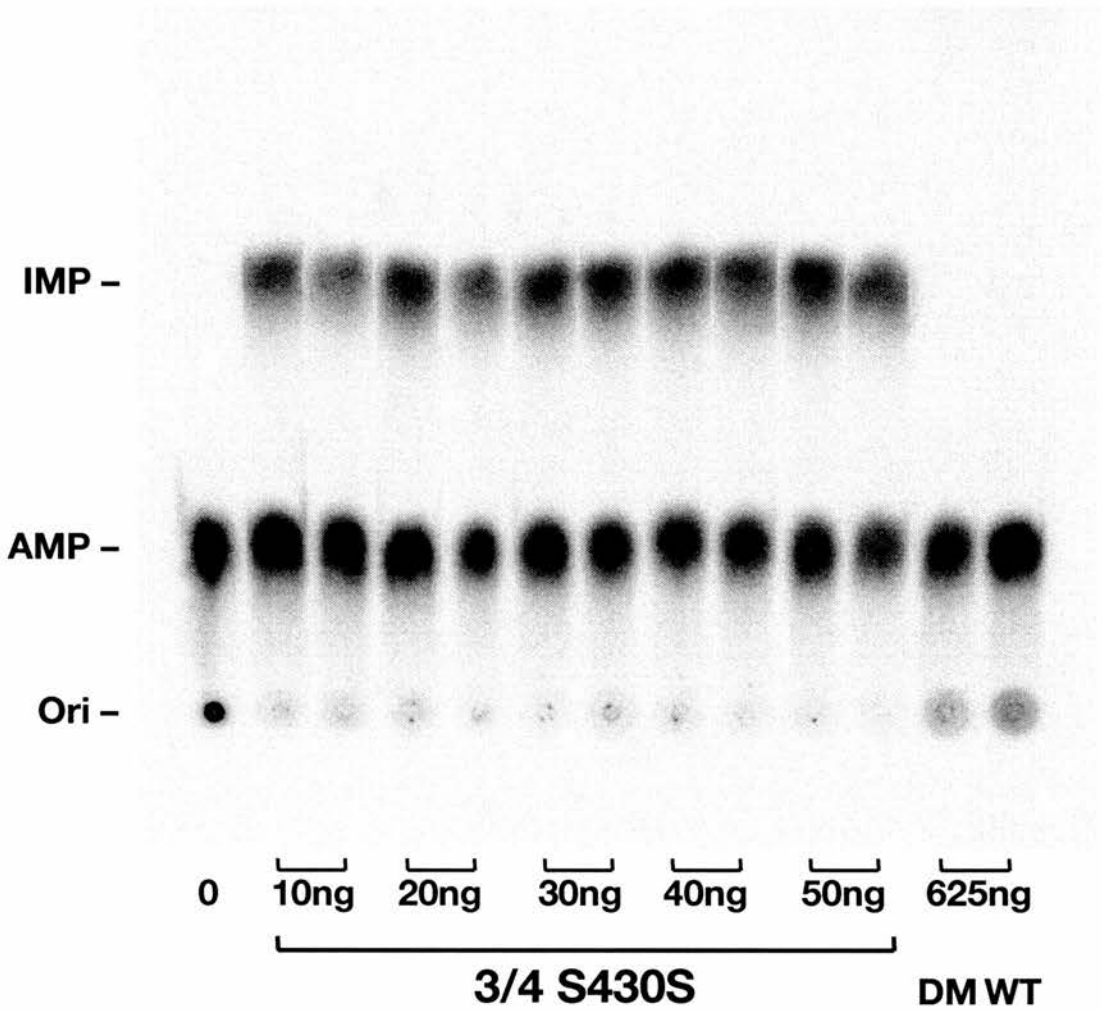
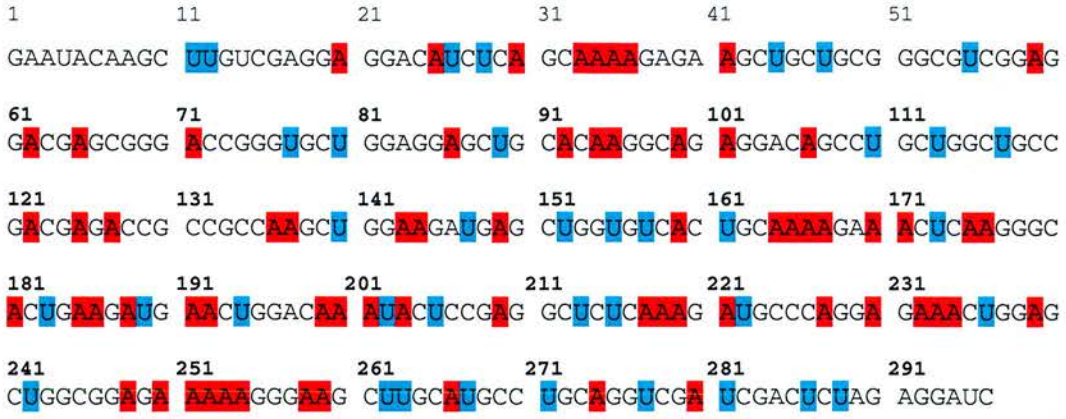

IMP –
AMP –
Ori –
0
10ng
20ng
30ng
40ng
50ng
625ng
3/4 S430S
DM WT

Figure 3.10A Editing assay comparing editing *in vitro* on a non-specific dsRNA substrate by dADAR 3/4 S430S and DM WT

3/4 S430S and DM WT were incubated with dsRNA labelled with α-P^{32}-ATP, the RNA digested, the products separated by thin layer chromatography and editing activity was quantified using a phosphoimager and ImageQuant™ software. The fastest migrating spots were IMP; inosine monophosphate, AMP; adenosine monophosphate, and the origin; Ori is at the bottom. The first lane contains no ADAR protein. Lanes 2 to 11 contain increasing amounts (125 pico moles to 625 pico moles) of 3/4 S430S and corresponds to editing of the dsRNA substrate to 28% (10 ng/125 pico moles), 37% (20 ng/250 pico moles), 42% (30 ng/375 pico moles), 44% (40 ng/500 pico moles) and 48% (50 ng/625 pico moles). Lanes 12 and 13 contain 13.75 nano moles DM WT and show no editing activity on the dsRNA substrate.

Figure 3.10Bi Mapping of dADAR 3/4 S430S editing on the ΔKP substrate

The sequence of the ΔKP substrate is shown at the top of the page. Adenosines highlighted in red indicate sites where 3/4 S430S has deaminated an adenosine and resulted in an A-to-G change on the sense strand. Uridines highlighted in blue indicate U-to-C changes on the anti-sense strand. Each editing site is listed and the percentage editing by 3/4 S430S calculated.



RED: Sense strand (A > G) **BLUE: Anti-sense strand (T > C)**

U	C/total no. seq	U	C/total no. seq	U	C/total no. seq
11	4/33 **(12%)**	140	34/41 **(83%)**	215	34/41 **(84%)**
12	4/33 **(12%)**	147	21/41 **(51%)**	222	4/41 **(9.75%)**
26	3/33 **(9%)**	152	39/41 **(95%)**	236	35/41 **(85%)**
28	27/33 **(81%)**	155	14/41 **(34%)**	242	26/40 **(65%)**
44	37/38 **(97%)**	157	6/41 **(15%)**	262	31/41 **(76%)**
47	37/38 **(97%)**	161	32/40 **(80%)**	263	11/41 **(27%)**
55	28/39 **(72%)**	173	4/40 **(10%)**	267	31/41 **(76%)**
77	36/41 **(88%)**	183	37/40 **(92.5%)**	271	35/41 **(85%)**
80	35/41 **(85%)**	189	3/40 **(7.5%)**	277	17/41 **(41%)**
89	38/40 **(95%)**	194	32/40 **(80%)**	281	8/41 **(19.5%)**
110	32/40 **(80%)**	202	39/40 **(97.5%)**	286	2/41 **(4.8%)**
113	34/40 **(85%)**	205	1/40 **(2.5%)**	288	37/41 **(90%)**
117	40/41 **(98%)**	213	17/41 **(42%)**		

A	G/total no. seq	A	G/total no. seq	A	G/total no. seq
20	1/16 **(6.25%)**	137	21/21 **(100%)**	209	9/23 **(39%)**
25	3/16 **(18.75%)**	143	5/21 **(24%)**	217	18/23 **(78%)**
30	7/16 **(43.75%)**	144	21/21 **(100%)**	218	21/23 **(91%)**
33	4/16 **(25%)**	149	12/21 **(57%)**	219	20/23 **(87%)**
34	12/16 **(75%)**	159	6/20 **(30%)**	221	4/23 **(17%)**
35	12/16 **(75%)**	164	16/21 **(76%)**	227	22/24 **(92%)**
36	13/17 **(76%)**	165	19/21 **(90%)**	230	8/24 **(33%)**
41	13/17 **(76%)**	166	19/21 **(90%)**	232	3/24 **(12.5%)**
59	1/18 **(5.5%)**	167	21/21 **(100%)**	233	23/24 **(96%)**
62	6/20 **(30%)**	170	19/22 **(86%)**	234	23/24 **(96%)**
65	13/20 **(65%)**	171	14/22 **(63%)**	239	4/24 **(17%)**
71	6/21 **(28%)**	175	5/22 **(22%)**	248	1/24 **(4%)**
86	15/21 **(71%)**	176	20/22 **(91%)**	250	1/24 **(4%)**
92	10/18 **(55%)**	181	18/22 **(81%)**	251	20/24 **(83%)**
94	3/18 **(16.5%)**	185	6/22 **(27%)**	252	20/24 **(83%)**
95	18/18 **(100%)**	186	20/22 **(91%)**	253	23/24 **(96%)**
99	14/19 **(74%)**	188	1/22 **(4.5%)**	254	22/24 **(92%)**
101	2/19 **(10.5%)**	191	2/23 **(8.7%)**	258	9/24 **(37.5%)**
106	14/19 **(74%)**	192	12/23 **(52%)**	259	23/24 **(96%)**
122	7/20 **(35%)**	199	4/23 **(17%)**	266	16/24 **(67%)**
125	16/20 **(80%)**	200	20/23 **(87%)**	274	20/24 **(83%)**
127	3/20 **(15%)**	201	20/23 **(87%)**	280	11/24 **(46%)**
136	17/21 **(81%)**	203	21/23 **(91%)**		

Figure 3.10Bii Mapping of dADAR DM WT editing on the ΔKP substrate

The sequence of the ΔKP substrate is shown at the top of the page. Adenosines highlighted in red indicate sites where DM WT has deaminated an adenosine and resulted in an A-to-G change on the sense strand. Uridines highlighted in blue indicate U-to-C changes on the anti-sense strand. Each editing site is listed and the percentage editing by DM WT calculated.

```
1           11          21          31          41          51
GAAUACAAGC UUGUCGAGGA GGACAUCUCA GCAAAAGAGA AGCUGCUGCG GGCGUCGGAG

61          71          81          91          101         111
GACGAGCGGG ACCGGGUGCU GGAGGAGCUG CACAAGGCAG AGGACAGCCU GCUGGCUGCC

121         131         141         151         161         171
GACGAGACCG CCGCCAAGCU GGAAGAUGAG CUGGUGUCAC UGCAAAAGAA ACUCAAGGGC

181         191         201         211         221         231
ACUGAAGAUG AACUGGACAA AUACUCCGAG GCUCUCAAAG AUGCCCAGGA GAAACUGGAG

241         251         261         271         281         291
CUGGCGGAGA AAAAGGGAAG CUUGCAUGCC UGCAGGUCGA UCGACUCUAG AGGAUC
```

RED: Sense strand (A > G) BLUE: Anti-sense strand (T > C)

A	G/total no. seq	T	C/total no. seq
17	8/48 **(16%)**	11	2/46 **(4.3%)**
30	1/49 **(2%)**	12	1/46 **(2.2%)**
59	2/49 **(4%)**	26	1/45 **(2.22%)**
94	1/47 **(2.1%)**	47	1/47 **(2.1%)**
101	1/47 **(2.1%)**	89	1/47 **(2.1%)**
104	1/47 **(2.1%)**	110	1/47 **(2.1%)**
144	2/47 **(4.25%)**	117	3/47 **(6.4%)**
167	1/47 **(2.1%)**	140	1/47 **(2.1%)**
188	1/47 **(2.1%)**	147	1/47 **(2.1%)**
191	1/47 **(2.1%)**	155	2/47 **(4.25%)**
200	1/47 **(2.1%)**	183	1/46 **(2.2%)**
201	1/47 **(2.1%)**	189	1/46 **(2.2%)**
254	1/47 **(2.1%)**	202	2/46 **(4.35%)**
259	1/47 **(2.1%)**	277	2/46 **(4.35%)**
280	19/36 **(53%)**	287	1/46 **(2.2%)**
284	2/35 **(5.8%)**		
289	12/30 **(40%)**		

sites than 3/4 S430S but those it did deaminate were also edited by the full length protein.

3.9 A SERINE TO GLYCINE MUTATION REDUCES hADAR2 EDITING ACTIVITY *IN VITRO*

Having demonstrated that the deaminase domain has activity and can edit its own transcript, I decided to further investigate the role of this editing event. It has been shown that the serine to glycine change within the deaminase domain of full length dADAR lowers editing activity (Keegan et al., 2005; Palladino et al., 2000a). However, does this editing event occur in other *adar* transcripts and does it have the same effect?

Sequence analysis of this region in other *adar* enzymes shows *Drosophila pseudoobscura, Anopheles* and *Culicoides* also have an AGT serine codon in this position therefore the possibility of editing at this site exists (see Figure 3.11). The human *adar2* gene also codes for serine at this position though in the form of an uneditable TCA codon. To assess whether the reduction in ADAR activity seen in the edited form of dADAR is conserved I decided to mutate the serine codon in *adar2* to glycine and compare its activity to the wild-type hADAR2 protein.

Point mutations in *adar2* were generated by oligonucleotide-directed mutagenesis (See section 2.1.6.3) to change the TCA serine codon to the GGA glycine codon. Mutagenesis was performed on *adar2* cloned in pPICZa, a *Pichia pastoris* vector, using the QuikChange™ site-directed mutagenesis kit (Stratagene) and the sequences analysed.

Wild-type hADAR2 and hADAR2 S/G were expressed and purified as described in section 2.2.1 and their ability to edit a non-specific dsRNA substrate assayed and the results analysed by thin layer chromatography.

```
trADAR2A        ----------------------------------TGTGAAAAGAAACAAGGT-TTCAGAC 25
trADAR2B        ----------GACCACCAGAAGTCAATATTTGTCCACTGTGAGAAGGGAGGC-TACAGGT 49
hADAR2          TAACAAAGATGATCAAAAAGAATCCATCTTTCAGAAATCAGAGCGAGGGGGG-TTTAGGC 59
dADAR           --CCACAGCATATCAGTCGATTTTCGTGAGGAATACTGATGGGCAATACCCT-TATAAAC 57
D.pseudo        --CCACAGCATATCAGTCGATTTTCGTGAGGAATACTGATGGGCAATACCCT-TATAAAC 57
csRED1          ----AGTGCAAAACAGTCAATTTTCCTTAA-ACCAGAAAATGGTAGCACAAAATATCGCT 55
A.gambiae       CAAGCGGGCACGGAAATC-ATTTTTAACACGCCCAAAGACGGTTCGAACCTG-TACACGC 58
                                                                        *

trADAR2A        TGAAGGAGAACGTCCAGTTCCACCTGTACATTAGCACCTCACCCTGCGGAGATGCTCGCA 85
trADAR2B        TAAAGGACAACGTACAATTTCACCTCTACATCAGCACTTCGCCCTGTGGAGACGCCAGGA 109
hADAR2          TGAAGGAGAATGTCCAGTTTCATCTGTACATCAGCACCTCTCCCTGTGGAGATGCCAGAA 119
dADAR           TAAAATCCGGGGTACATTTCCATTTGTATATAAATACAGCACCTTGTGGGGATGCACGGA 117
D.pseudo        TAAAATCCGGGGTACATTTCCATTTGTATATAAATACAGCACCTTGTGGGGATGCACGGA 117
csRED1          TGCGACCCGAAATTCAATTCCATTTGTATATAAATACAGCGCCGTGTGGTGATGCTCGAG 115
A.gambiae       TGAAGGACGGCATCTACTTCCATCTGTACATCAACACGGCACCGTGCGGCGATGCGCGCG 118
                *          *  * ** **  * ** ** *  **  * ** ** ** ** **  *
                      Ser
trADAR2A        TCTTCTCTCCTCACGAGGCTGGAGTGGAAGATCAGGGAGACAGGCACCCTAAC-- 138
trADAR2B        TCTTCTCTCCACACGAAGCCGGCGTTGAAGATCAGGGAGACAGACACCCCAAC-- 162
hADAR2          TCTTCTCACCACATGAGCCAATCCTGGAAG------------------------- 149
dADAR           TATTTAGTCCTCACGAAAACGACACTGGT------GTTGATAAACATCCAAATAG 166
D.pseudo        TATTTAGTCCTCACGAAAACGACACAGGT------GTTGATAAACATCCAAATAG 166
csRED1          TATTTAGTCCGCATGAAGCTGACACAATTAAC---GGTGATAAGCATCCAAAT-- 165
A.gambiae       TCTTCAGCCCGCACGAAAACGACAGCATGGAC---GTGGACAAGCATCCGAAT-168
                * **    ** ** **
```

Figure 3.11 Alignment of *adar* gene sequences corresponding to the RNA editing substrate *exon 7* in the *Drosophila melanogaster* deaminase domain.

The ADAR sequences are shown aligned by the Clustal W multiple sequence alignment program (http://www.ebi.ac.uk/clustalw/) with the corresponding regions of *Takifugu rubripes* ADAR2a and ADAR2b, human ADAR2 (exon 5), *Drosophila melanogaster* ADAR (exon 7), *Drosophila pseudoobscura* ADAR (exon 7), *Culicoides sonorensis* ADAR2 and *Anopheles gambiae*. Conserved residues are highlighted by an *. The serine codon of dADAR modified by RNA editing to glycine is highlighted in red. Gaps are shown by dashed lines.

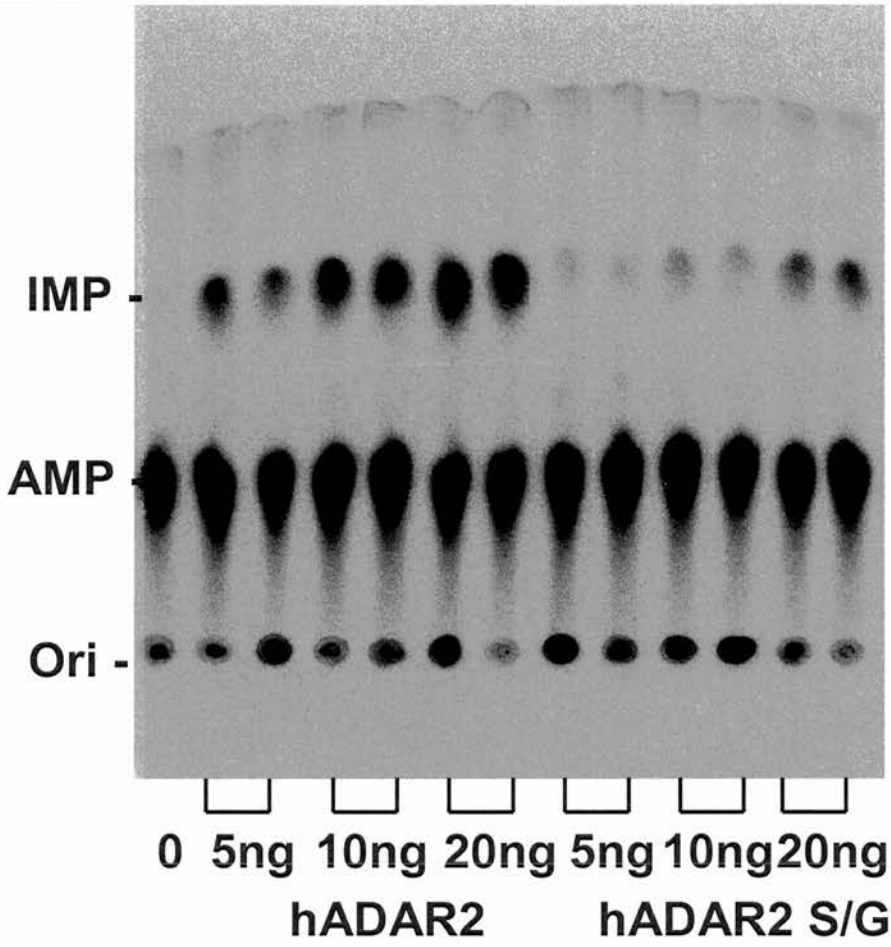


Figure 3.12 Editing assay comparing editing *in vitro* on a non-specific dsRNA substrate by hADAR2 and hADAR2 S/G.
hADAR2 and hADAR2 S/G were incubated with dsRNA labelled with α-^{32}P-ATP, the RNA digested and the products separated by thin layer chromatography. Editing activity was quantified using a phosphoimager and ImageQuant™ software. The fastest migrating spots were IMP; inosine monophosphate, AMP; adenosine monophosphate, and the origin; Ori is at the bottom. The first lane contains no ADAR protein. Lanes 2 to 7 contain increasing amounts of hADAR2 and corresponds to editing of the dsRNA substrate to 5.8% (5 ng/72.5 pico moles), 11% (10 ng/125 pico moles), 20% (20 ng/250 pico moles). Lanes 8 to 13 contain hADAR2 S/G and corresponds to editing of the dsRNA substrate to 0.6%% (5 ng/72.5 pico moles), 1.4% (10 ng/125 pico moles), 4.6% (20 ng/250 pico moles).

The editing activity of the mutated hADAR2 S/G protein on this substrate was around 5 - fold less that of the wild-type protein (Figure 3.12).

I now decided to assay the activity of the wild-type and mutant hADAR2 proteins on the mini B13 GluR-B Q/R transcript. As with the non-specific dsRNA substrates, hADAR2 S/G was less active when compared to the wild-type hADAR2. This time however, the reduction in activity was around 8 – fold (Figure 3.13A and 3.13B).

These results indicate that this serine residue influences ADAR activity however, the role it plays is unclear. A number of residues were identified that co-ordinate with IP_6 but serine 458 does not appear to be involved. One hypothesis suggests that the serine residue is phosphorylated and its change to glycine prevents this phosphorylation and thus affects enzyme activity however, this has yet to be shown.

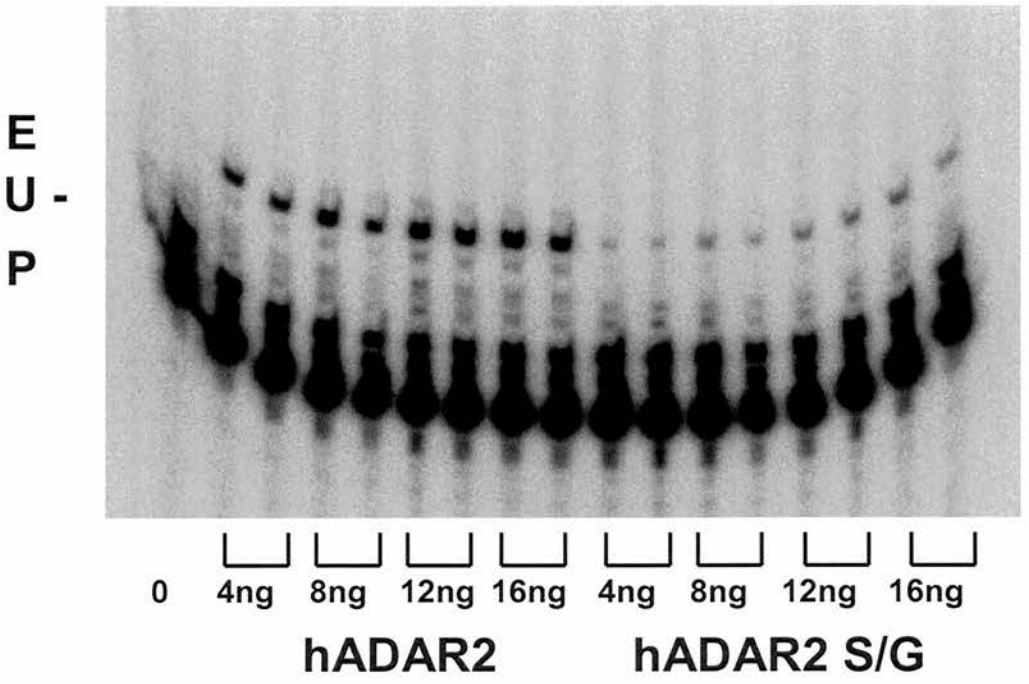


Figure 3.13A Poisoned primer extension assay comparing editing *in vitro* on *mini B13 GluR-B Q/R* site by hADAR2 and hADAR2 S/G. hADAR2 and hADAR2 S/G were incubated with 10 femtomoles *mini B13 GluR-B Q/R* transcript for 1 hour. The first lane contains no ADAR protein. Lanes 2 to 9 contain increasing amounts of hADAR2 and corresponds to editing of the *mini B13* substrate to 17% (4 ng/50 pico moles), 28% (8 ng/100 pico moles), 34.5% (12 ng/150 pico moles) and 36% (16 ng/200 pico moles). Lanes 10 to 17 contain increasing amounts of hADAR2 S/G and corresponds to editing of the *mini B13* transcript to 2% (4 ng/50 pico moles), 4% (8 ng/100 pico moles), 6% (12 ng/150 pico moles) and 6.4% (16 ng/200 pico moles). Editing activity was quantified using a phosphoimager and ImageQuant™software. Radiolabelled primer; P, the unedited; U and edited; E primer extensions products.

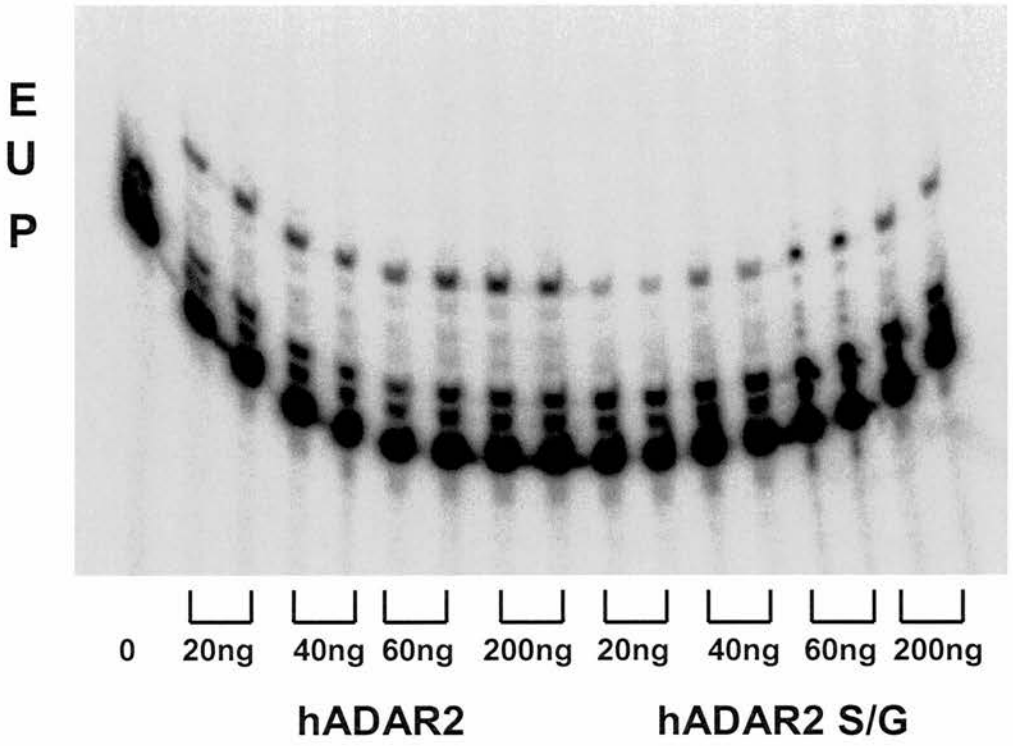


Figure 3.13B Poisoned primer extension assay comparing editing *in vitro* on *mini B13 GluR-B Q/R* site by hADAR2 and hADAR2 S/G. hADAR2 and hADAR2 S/G were incubated with 10 femto moles *mini B13 GluR-B Q/R* transcript for 1 hour. The first lane contains no ADAR protein. Lanes 2 to 9 contain increasing amounts of hADAR2 and corresponds to editing of the *mini B13* substrate to 31% (20 ng/250 pico moles), 32.5% (40 ng/500 pico moles), 35% (60 ng/750 pico moles) and 35% (200 ng/2.5 nano moles). Lanes 10 to 17 contain increasing amounts of hADAR2 S/G and corresponds to editing of the *mini B13* transcript to 8.6% (20 ng/250 pico moles), 13% (40 ng/500 pico moles), 14% (60 ng/750 pico moles) and 17% (200 ng/2.5 nano moles). Editing activity was quantified using a phosphoimager and ImageQuant™software. Radiolabelled primer; P, the unedited; U and edited; E primer extensions products.

3.10 Discussion

The catalytic domain of the *Drosophila* ADAR enzyme is highly conserved sharing active site motifs with the family of cytidine deaminases. The function of these motifs has been well studied and the recent resolution of the crystal structure of the deaminase domain of hADAR2 has documented the individual roles assigned to these conserved residues (Macbeth et al., 2005). Initially a series of deaminase constructs were designed including an edited isoform, an unedited isoform, an inactive isoform and the wild-type domain. However, these appeared to be unstable on expression resulting in the purification of a *Pichia* contaminant. A new construct was designed extending the amino terminus to the 3′ end of the second dsRBM and this was expressed successfully. This construct is very similar to the hADAR2-D construct crystallized by Macbeth *et al.*, with only 4 residues different at the amino terminus (Macbeth et al., 2005) (see appendix 1).

Both dADAR and hADAR2 are very similar proteins functionally as human *adar2* is able to rescue null mutant *Adar* flies (Keegan, unpublished). Structurally they share the same tertiary pattern with both proteins consisting of two dsRNA binding domains and a catalytic domain unlike human ADAR1 which consists of three dsRBDs and has an extended amino terminus containing Z-DNA binding domains. However, a number of post transcriptional modifications occur in dADAR that influence the proteins activity which are not seen in either human ADAR1 or ADAR2.

In particular, the serine to glycine change that occurs in the deaminase domain of dADAR has been shown to reduce enzymatic activity (Keegan et al., 2005). This self editing is believed to be a method of negative auto regulation and has so far only been detected in the *D. melanogaster* protein. Sequence analysis of other ADAR proteins show the ADAR2-like proteins have serine in this position and as my results demonstrated, the mutation of

this serine to glycine in another ADAR protein, in this case human ADAR2, has the same effect and decreases editing activity.

The functional cause of this decrease in activity is unclear as a specific role has not yet been assigned to this residue, though as discussed it has been hypothesised that it may be phosphorylated. Analysis of the dADAR full length protein did highlight a number of phosphorylated residues across the protein however this serine was not one of them (M. O'Connell, unpublished).

Interestingly ADAR1-like proteins have a negatively charged aspartic acid residue in this position (See figure 3.1). When the serine residue of dADAR is mutated to aspartic acid the protein is inactive (M. O'Connell, unpublished) indicating that not only do they have different numbers of dsRBDs, the deaminase domains of dADAR and hADAR2 differ from that of hADAR1.

One common characterisitic they share though is the ability of their catalytic domains to function without the dsRBDs. I have demonstrated that the deaminase domain of dADAR is able to edit *exon 7* transcripts *in vivo* and non-specific dsRNA *in vitro* though at much lower levels than the full length protein. This corresponds to results from Herbert *et al.* and Macbeth *et al.* who demonstrated *in vivo* ADAR1 and *in vitro* ADAR2 deaminase domain activity, respectively (Herbert and Rich, 2001; Macbeth et al., 2005).

Macbeth was able to demonstrate hADAR2 deaminase domain activity on non-specific dsRNA using thin layer chromatography unlike my own results on the *Drosophila* ADAR deaminase domain which required the more sensitive RT-PCR to detect any editing activity. As discussed earlier, the constructs were very similar therefore it would seem unlikely that a structural difference between the two construct designs could account for the different editing activity.

As I discussed in section 3.7, the editing activity of the full length dADAR control is much lower than normally seen so perhaps some DM WT activity could be detected on a more efficiently edited sample of *exon 7*. However, the difference in activity on the non-specific dsRNA substrate is less easily explained and may simply be due to inherent differences between the human and *Drosophila* proteins.

The fact that the human and *Drosophila* deaminase domains are able to edit transcripts at all is surprising. A number of studies have demonstrated that both proteins function as a homodimer and that dimerisation is required for activity (Gallo et al., 2003; Jaikaran et al., 2002). A distinct region of interaction for hADAR2 has yet to be identified however; Gallo *et al.* clearly showed that the amino terminus and first dsRBM were required for the dimerisation of dADAR and hence protein activity. My results would suggest that this is not the the whole story. The recruitment of dADAR monomers onto the dsRNA substrate and the mechanism by which the deamination reaction occurs is unclear but the possibility of even weak interactions between points in the deaminase domain would make sense.

Mapping the editing sites of both full length and DM WT on *exon 7* and the non-specific dsRNA ΔKP illustrated two things. Firstly, regarding the non-specific dsRNA, the full length protein will repeatedly edit the same adenosine bases indicating that not all adenosines along a dsRNA substrate are suitably positioned for editing and that even within a duplex there is a basal level of specificity. Secondly, the pattern of editing seen with the full length protein is repeated at a number of sites for the deaminase domain even though efficiency is greatly reduced. This specificity is more clearly seen when analysing the *exon 7* substrate as the main difference here is not in site selection but in editing efficiency. The suggestion that the deaminase domain plays a role in site selection is supported by Wong *et al.* who clearly demonstrated that the deaminase domain of hADAR1 and hADAR2 retained editing site specificity (Wong et al., 2001).

More recently, the editing patterns of hADAR1, hADAR2 and dADAR on a non-specific substrate have been analysed. As expected, all three proteins exhibited distinct but overlapping specifities with the *Drosophila* protein editing many sites more efficiently than the human proteins (Scadden and O'Connell, 2005).

Overall the structural and functional differences between each of these proteins may influence the obvious difference in specificity they have for both physiological transcripts and adenosines along a non-specific dsRNA duplex.

4

THE *DROSOPHILA* ADAR dsRNA BINDING DOMAIN

4.1 INTRODUCTION

The dsRNA binding motif (dsRBM) was first identified in the *Xenopus laevis* RNA-binding protein A (XlrbpA) and the *Drosophila* maternal effect protein Staufen (St Johnston et al., 1992). It has since been identified in a number of diverse proteins involved in the regulation and modification of RNA molecules. The ADAR family of proteins contain two or more dsRNA binding motifs tandemly arranged at the amino terminus. Studies have shown that not all of the dsRBMs are neccessary for enzyme activity but one is normally essential (Lai et al., 1995; Liu and Samuel, 1996).

I have cloned and purified the *Drosophila* dsRNA binding domains and analysed their structural and binding characteristics. The results suggest that they are *bona fide* members of the dsRBM family. The co-expression of these domains alongside the catalytic domain *in vivo* alters the editing activity of the deaminase domain indicating that the dsRBMs can indeed interact with dsRNA substrates and facilitate the editing reaction.

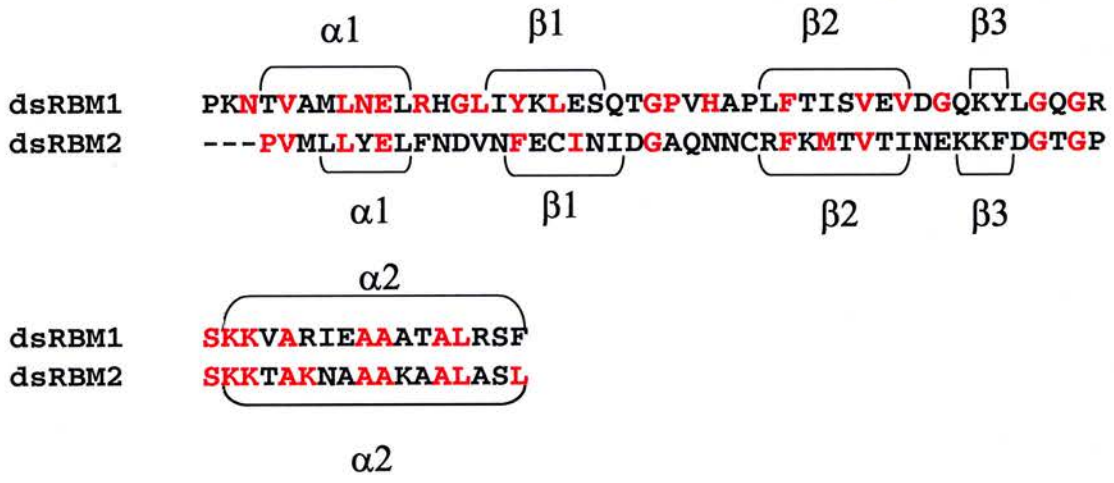


Figure 4.1 Alignment of the two dsRNA binding motifs found in *Drosophila* ADAR. Amino acid sequences of the two dsRBMs are given in single-letter code. Letters coloured red indicate amino acids found in the dsRBM consensus sequence. The α-helical and β-sheet regions predicted by primary sequence analysis are indicated by the brackets.

4.2 Sequence and Structural analysis of the dsRBMs

Analysis of the *D. melanogaster* ADAR dsRNA binding motifs dsRBM1 and dsRBM2 demonstrates that they are 63 and 60 amino acids in length, respectively and are share a similar sequence. They show strong homology with the dsRBM consensus sequence (Doyle and Jantsch, 2002) across their full length indicating they are type A dsRBMs and will bind dsRNA strongly (See figure 4.1). Indeed, targeted point mutations changing alanine 106 and 110 to glutamate in the first dsRNA binding motif have been shown to abolish dsRNA binding activity of the full length protein (Gallo et al., 2003).

Secondary structure prediction programmes suggest a typical αββββα structure. The HSQC spectrum of dsRBM1 indicated the protein was well folded (See figure 4.2). Data from two-dimensional and three-dimensional NOESY provided information on interproton distances whilst two dimensional TOCSY and COSY allowed the assignments of amino acid side chains (Data provided by Jose Gallego, MRC LMB). Taken together these data show that dsRBM1 is indeed a typical dsRNA binding motif with a α_1-β–β–β–α_2 structure.

4.3 Expression and purification of dADAR dsRNA binding domain

An alternatively spliced exon, 3a, is located between the two dsRNA binding motifs of dADAR and the presence or absence of this exon alters the spacing between these two motifs (Palladino et al., 2000a). This has been shown to

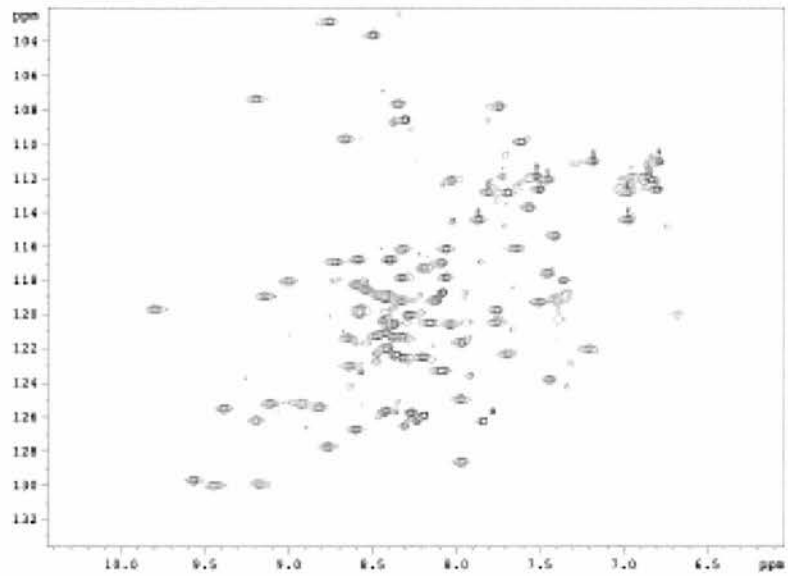


Figure 4.2 ^{15}N-^{1}H HSQC spectrum of *Drosophila* ADAR dsRBM1. Each signal in the HSQC spectrum represents a proton that is bound to a nitrogen atom. Since there is only one backbone H^N per amino acid, each HSQC signal represents one single amino acid. The HSQC also contains signals from the NH_2 groups of the side chains of Asn and Gln and of the aromatic H^N protons of Trp and His. Analysis of RBM1 with the 6x His tag cleaved results in a good dispersion pattern and indicates a well folded protein. Data provided by Jose Gallego.

affect the activity of the dADAR protein with exon 3a containing transcripts less active than 3/4 isoforms.

A number of constructs were designed to allow analysis of these dsRNA binding domains together with and without the alternatively splice exon as well as separately (See figure 4.3A). Constructs were cloned into the pET-28a vector in frame with a 6x HIS tag and expression in *E. coli* was under the control of the *T7lac* promoter. To assess the efficiency of induction culture samples were taken prior to induction and at hourly intervals after the addition of IPTG. Analysis by SDS-PAGE and Coomassie staining showed all proteins were expressed except the full domain consisting of both dsRBMs and exon 3a (See figure 4.3B). Further analysis indicated that this protein was cleaved. The presence of exon 3a in full length dADAR lengthens the relatively unstructured linker region between the two dsRBMs probably increasing the overall flexibility of the domain. When expressing a recombinant version of the domain alone this unstructured linker region will probably be more susceptible to proteolytic cleavage.

The His-tagged RBM proteins were initially purified over Ni^{2+}-NTA agarose in 500 mM KCl. SDS-PAGE analysis detected bands corresponding to the predicted molecular weight of the constructs however a number of contaminant proteins were also present. Further elution over a second Ni^{2+}-NTA column at 500 mM KCl followed by a 1 M KCl column improved the purification though an *E. coli* contaminant of approximately 45 kDa remained present especially in the purification of RBM1 +RBM2 minus exon 3a. (See figures 4.4A, 4.4B and 4.4C)

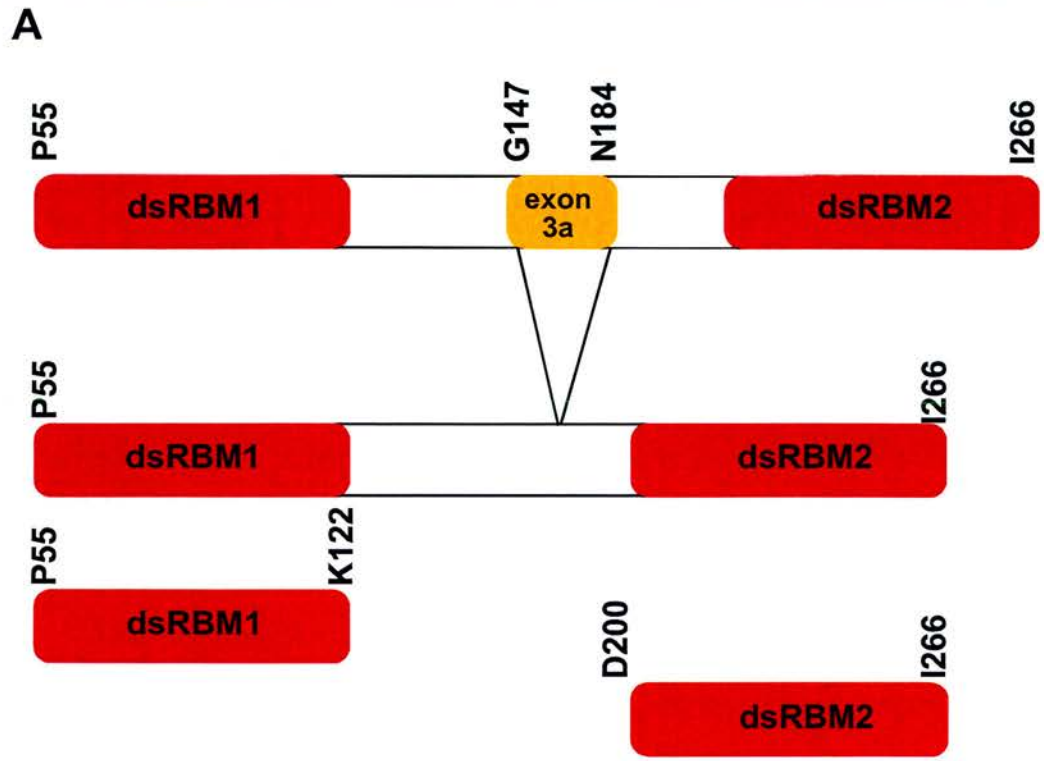


Figure 4.3A Schematic of the dsRBM constructs expressed in *E. coli*. Constructs containing RBM1 were designed to begin translation at proline residue 55 of the full length dADAR protein and a C-terminal tail containing the 6x histidine tag was added. The RBM2 construct begins at aspartate residue 200.

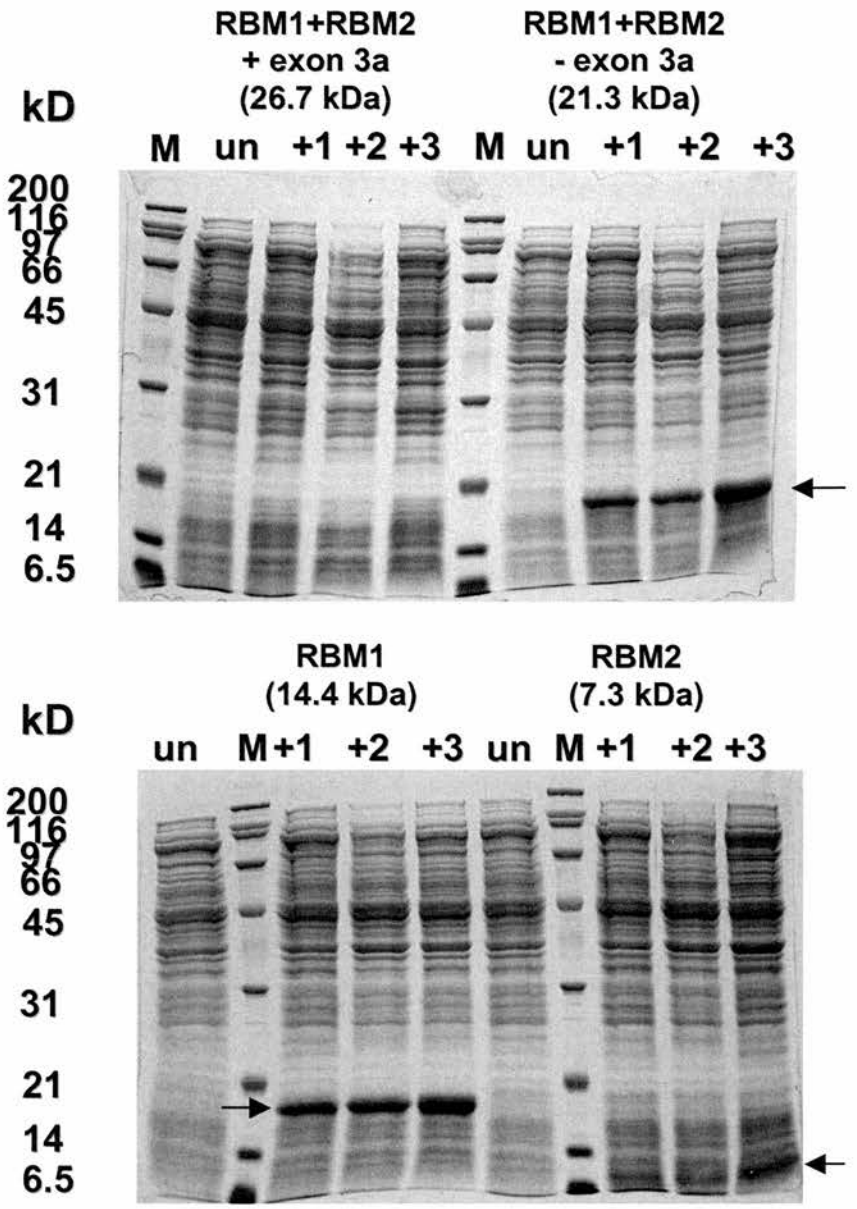


Figure 4.3B The induction and expression of the dsRBM constructs in *E. coli*. Expression of His-tagged RBM proteins in *E. coli* was induced by IPTG. Culture samples were taken prior to induction (un) and at 1 hour, 2 hours and 3 hours after IPTG induction. The samples were analysed by SDS-PAGE and Coomassie staining. Arrows indicate the expressed proteins. No expression was seen for the construct containing both dsRBMs and exon 3a. Protein markers (kDa) are indicated on the left.

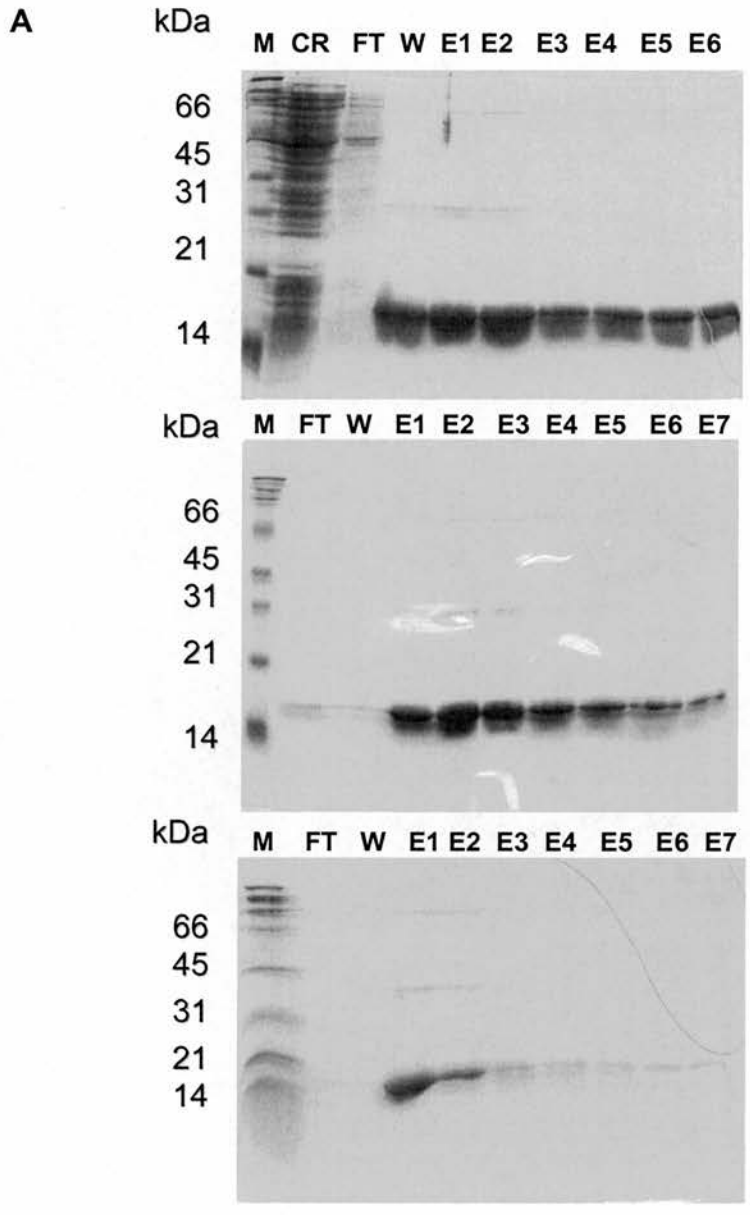


Figure 4.4A The purification of RBM1

His-tagged RBM1 protein was purified from *E. coli* over a series of Ni^{2+}NTA columns initially at a stringency of 500 mM KCl. This was then repeated with the pooled and dialysed elutates followed by a third column at 1 M KCl for the final purification. Samples were taken from the crude (CR), flow through (FT), wash (W) and elutions (E1-E7) and the level of purification at each stage was determined by SDS-PAGE and Coomassie staining. Protein markers (kDa) are indicated on the left.

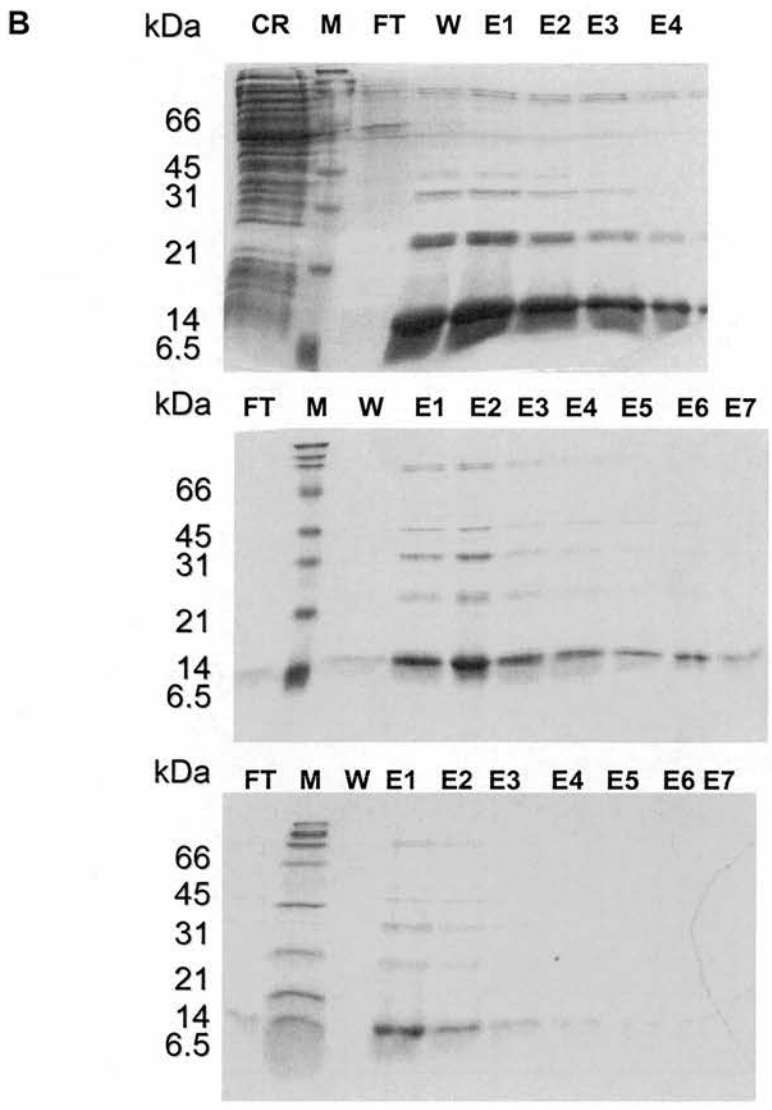


Figure 4.4B The purification of RBM2

His-tagged RBM2 protein was purified from *E. coli* over a series of Ni^{2+}NTA columns initially at a stringency of 500 mM KCl. This was then repeated with the pooled and dialysed elutates followed by a third column at 1 M KCl for the final purification. Samples were taken from the crude (CR), flow through (FT), wash (W) and elutions (E1-E7) and the level of purification at each stage was determined by SDS-PAGE and Coomassie staining. Protein markers (kDa) are indicated on the left.

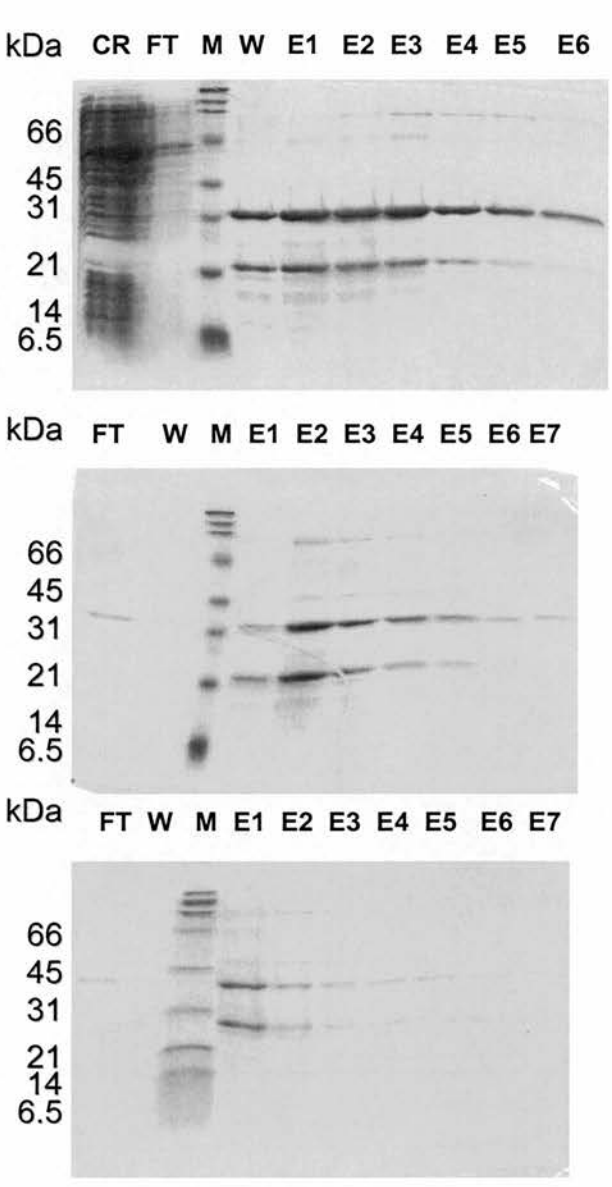


Figure 4.4C The purification of RBM1 + RBM2 minus exon 3a

His-tagged RBM1 + RBM2 minus exon 3a protein was purified from *E. coli* over a series of Ni^{2+}NTA columns initially at a stringency of 500 mM KCl. This was then repeated with the pooled and dialysed elutates followed by a third column at 1 M KCl for the final purification. Samples were taken from the crude (CR), flow through (FT), wash (W) and elutions (E1-E7) and the level of purification at each stage was determined by SDS-PAGE and Coomassie staining. Protein markers (kDa) are indicated on the left.

4.4 THE dADAR dsRNA BINDING MOTIFS CAN BIND dsRNA AND INFLUENCE RNA EDITING

To assess the binding activity of the dsRBMs filter binding assays were performed using radiolabelled ΔKP dsRNA as the ligand. A series of dilutions ranging from 31.25 nM to 24 μM were incubated with 200 fmoles of dsRNA for 10 min and filtered on a nitrocellulose membrane. The filters were washed twice, dried and the radioactivity measured in a scintillation counter in the presence of scintillation liquid. Two control samples of unfiltered dsRNA and a sample of washed dsRNA were also measured to assess the load and background values. The percentage of dsRNA bound was calculated and plotted against the protein amounts (See figure 4.5A). RBM1 saturated at around 4 μM. RBM2 failed to saturate even at levels of 24 μM however, circular dichroism and UV spectroscopy indicated nucleic acid contamination of the RBM2 protein (See figure 4.5B). This was repeated a further two times with no change in the result. Further purification steps were attempted including ion exchange chromatography to remove the nucleic acid contamination but these steps were unsuccessful. It therefore appears that this contamination may be RNA from *E. coli* binding to RBM2 and therefore preventing it binding to the radiolabelled dsRNA hence the reason saturation point was never reached.

4.5 THE dsRBMs CAN BE CO-EXPRESSED WITH DM WT IN *PICHIA PASTORIS*

A cloning strategy was devised to enable the dsRNA binding domain constructs to be co-expressed with DM WT in *Pichia pastoris*. This would allow me to assay whether the dsRBMs can influence the *in vivo* editing

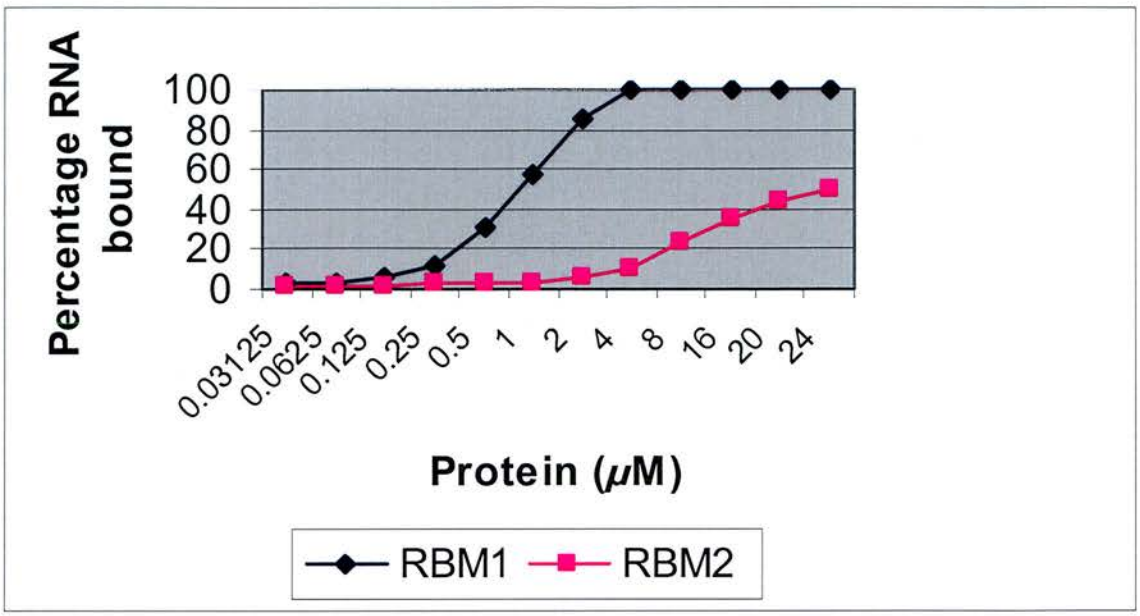


Figure 4.5A Filter binding of RBM1 and RBM2 to dsRNA. Filter binding was performed with the two dsRNA binding motifs; the protein amounts are indicated on the x-axis in μM.

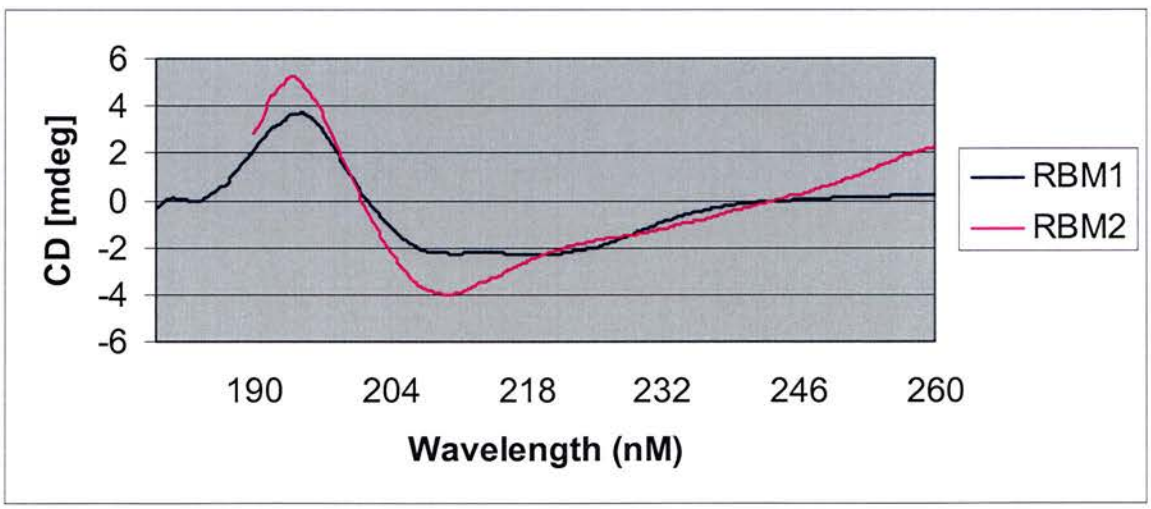


Figure 4.5B Far UV CD spectroscopy of RBM1 and RBM2. Far UV circular dichroism spectroscopy of RBM1 and RBM2 was performed. Nucleotide contamination is evident in RBM2 indicated by the continued rise in signal after 240 nM.

activity of DM WT on the *Adar exon 7* substrate. The DM WT construct cloned into pPICZ A was integrated into the *Pichia* KM71 genome at the *aox::ARG4* loci and clones were screened for Zeocin resistance. Meanwhile, the dsRBM constructs were cloned into pHIL-D2 and integrated into DM WT containing KM71 strains at the *his4* locus before being screened for the His^+ phenotype (See figure 4.6A). His^+ $Zeocin^R$ clones were selected and grown as before in glycerol media before protein expression was induced by transferring the cells into a methanol medium. Expression of each protein was assayed by western blot analysis using an anti-tetra HIS antibody. DM WT co-expressed well with RBM2, RBM1+2 plus exon 3a and RBM1+2 minus exon 3a. RBM1 was expressed though a longer exposure was required to view the band. As before, RT-PCR and sequence analysis was performed on total RNA isolated from *Pichia* expressing DM WT and each of the RBM constructs. The results are shown in Table 4.1 below.

Table 4.1

Protein	No. clones Sequenced	No. of adenosines at S/G site	Percentage editing
DM WT + RBM1	20	3	15%
DM WT + RBM2	21	9	42.8%
DM WT + RBMs 1+2 minus exon 3a	18	0	0%
DM WT + RBMs 1+2 plus exon 3a	28	2	7%
DM WT	43	4	9.3%
dADAR 3/4 WT	20	13	65%

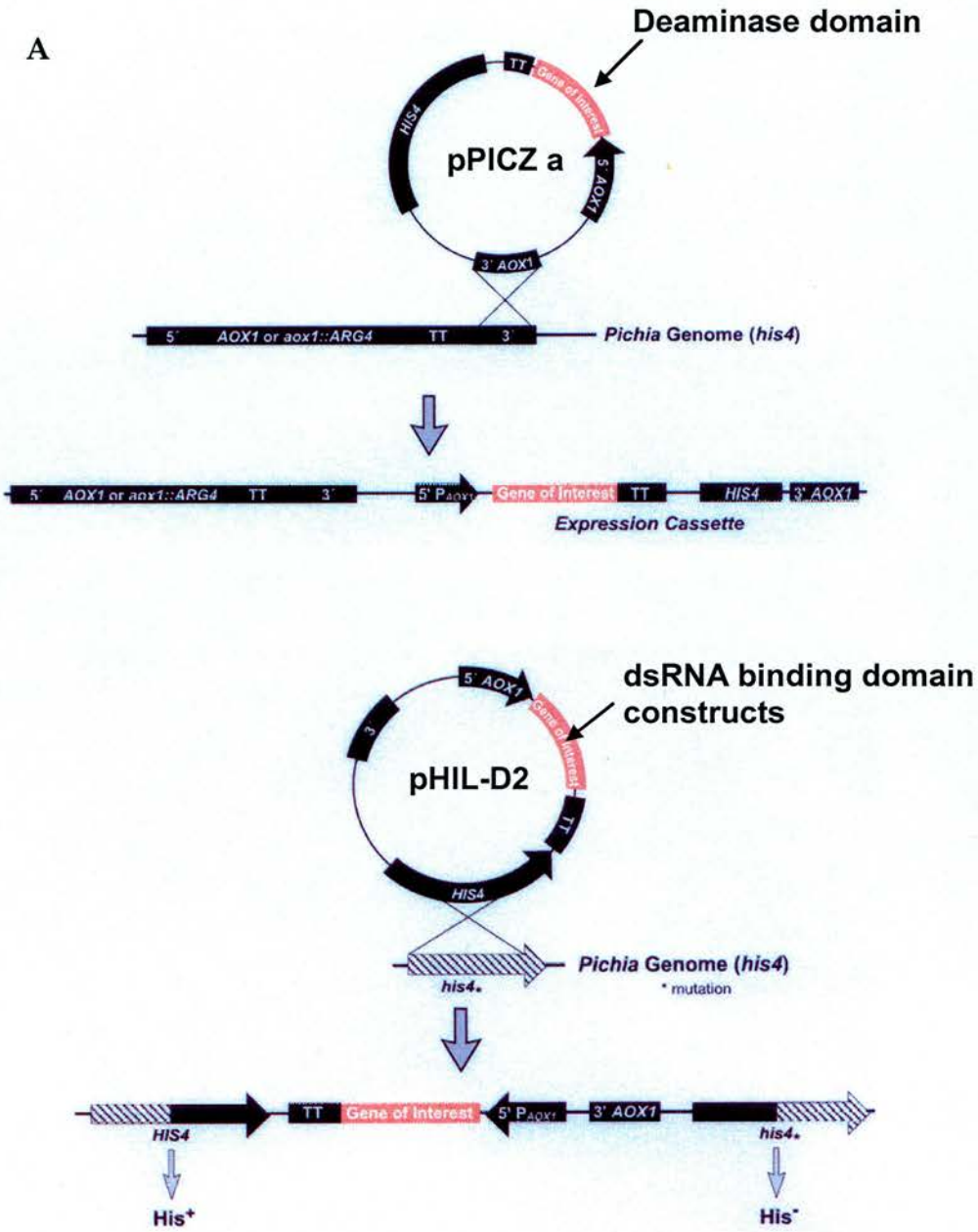


Figure 4.6A Schematic of the cloning strategy used for co-expressing the dsRBM constructs in *Pichia pastoris* with DM WT
The DM WT construct was cloned into pPICZ A and integrated into the *Pichia* KM71 strain genome at the *aox::ARG4* loci. Positive clones were isolated by Zeocin resistance screening. The dsRBM constructs were cloned into pHIL-D2 and integrated into the DM WT containing Zeocin resistant KM71 strains at the *his4* locus. Positive clones demonstrated a His^+ phenotype.

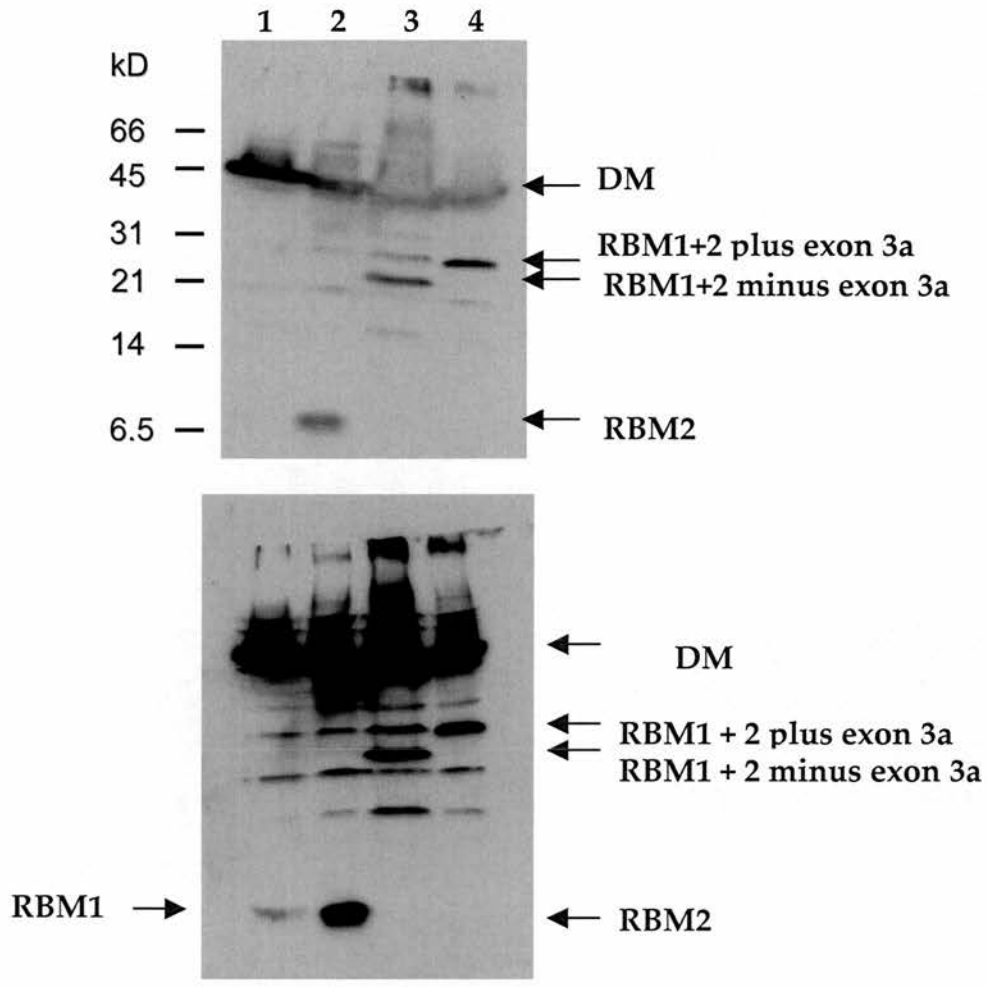


Figure 4.6B Co-expression of dsRBMs and DM WT in *Pichia pastoris*
After screening His^+ $Zeocin^R$ colonies were selected and grown as described. After induction samples were analysed by Western blot using an anti-tetra HIS antibody (Qiagen). Arrows indicate the expressed proteins. Lane 1 contains the DM WT plus RBM1 sample. Expression of RBM1 can be detected with longer exposures seen in the lower gel.

According to the data it appears RBM2 has the greatest effect on DM WT editing activity. It should be noted that it is impossible to control the relative amounts of protein expressed in *Pichia* as it is very dependent on the insertion locus and in the case of DM WT and integration in the *AOX1* region, multiple insertion events are possible. However, in this case the expression levels of DM WT should be constant as the same original DM WT KM71 colony was used for the integration of the RBMs. The comparison of the relative amounts of protein by western analysis is possible as the lanes were equally loaded therefore it appears that RBM1 has little effect due to its lower expression levels. Interestingly, the constructs consisting of both dsRBMs also had little effect on DM WT editing activity even though expression was detected. This could be due to construct design resulting in a incorrectly folded domain which interacts less well with dsRNA than the individual motifs.

4.6 Discussion

dsRNA binding domains consist of one or more dsRNA binding motifs tandemly arranged. In the case of dADAR two dsRNA binding domains are situated at the amino terminus. The hypothesis is that the two motifs and the linker region between them form a dumbell-like shape that wraps itself around a dsRNA substrate allowing the catalytic domain to deaminate an adenosine base.

Interestingly the linker region in dADAR separating these motifs contains an alternatively spliced exon, 3a, whose presence or absence changes the spacing and has been found to affect the overall activity of the protein (Palladino et al., 2000a). Isoforms with exon 3a are less active than those without 3a. It is believed that the extra residues between the motifs increase the flexibility of the whole domain and perhaps in the case of dADAR prevent it from binding as efficiently to a dsRNA substrate. Structural

analysis of this region should clarify the role of exon 3a and define the characteristics of both dsRNA binding motifs.

An attempt was made to express RBM1, RBM2 as well as both motifs with and without exon 3a in *E.coli*. As discussed earlier, I was unable to express the construct consisting of both motifs with exon 3a and believe this to be due to proteolytic cleavage within the unstructured linker region. However, successful expression of RBM1 allowed us to perform NMR and resolve the structure. As with other well documented dsRBD's (Bycroft et al., 1995; Kharrat et al., 1995; Nanduri et al., 1998), RBM1 revealed the common α–β–β–α structure in which the two α-helices lie on one face of a three-stranded anti-parallel β-sheet. Sequence analysis also indicated that both RBM1 and RBM2 were type A domains with the conserved residues at their carboxy terminus and would likely display dsRNA binding activity (Doyle and Jantsch, 2002).

The dsRBMs of a number of proteins have been found to perform different functions with some displaying no dsRNA binding activity but instead playing a role in nuclear localisation, protein dimerisation or activiation of transgene expression. Gallo *et al.* have already demonstrated that RBM1 is involved in protein dimerisation and that point mutations within this region can disrupt dsRNA binding activity preventing protein-protein interaction (Gallo et al., 2003). I performed filter binding analysis on recombinant RBM1 and RBM2 to compare the relative binding affinities of each motif. Both were able to bind dsRNA with high affinity however, the recombinant RBM2 samples appeared to be contaminated with dsRNA preventing a clear and accurate result. Further purification steps in up 1 M KCl did not change the result and RBM2 continued to bind unefficiently to dsRNA.

Having analysed both dsRBMs and the deaminase domain and found that all are able to function in isolation, I opted to co-express these constructs in *Pichia* and assay the editing activity of the deaminase domain on the *exon7*

transcript. As mentioned previously, the hADAR1 and hADAR2 deaminase domains have been shown to be active on their own and a number of deletion analyses of both proteins have identified the important regions essential for activity. However, no one has expressed the domains individually and then assayed editing activity.

My results showed that when co-expressed with the deaminase domain RBM2 can increase editing activity by around 4-fold. Many deletion analyses have been done on the human ADARs and all have concluded that the dsRNA binding motif immediately upstream of the deaminase domain is critical for catalytic activity (Lai et al., 1995; Liu and Samuel, 1996; Macbeth et al., 2004). My results would suggest the same as the editing activity of DM WT was greatly increased in the presence of RBM2.

Recent work by Stefl *et al.* clearly demonstrates that both RBMs of hADAR2 interact with the R/G substrate at different sites and at distinctive motifs on the dsRNA stem loop structure (Stefl et al., 2006). Each RBM contributes to the editing reaction in either a substrate-specific manner, like RBM1, or is essential for all editing, like dsRBM2 (R.B. Emeson, unpublished). Owing to their homology, the possibility exists that the dADAR RBM2 plays a similar role to that of the hADAR2 RBM2 and is essential for editing. This would explain the increase in RNA editing activity of the DM WT protein when co-expressed with RBM2.

More interesting though is that neither of the constructs containing both dsRBMs was able to improve editing efficiency, in fact if anything they appear to have a negative effect. As discussed by Gallo *et al.* the amino terminus and RBM1 of dADAR are essential for protein dimerisation and dADAR activity and removal of the amino terminus abolishes all activity (Gallo et al., 2003). This differs from hADAR2 where the amino terminus is not required for RNA editing (Macbeth et al., 2004) and is not involved in dimerisation (Poulsen, 2006). None of the RBM constructs I expressed

contained the amino terminus perhaps explaining the lack of co-operation between the domain and the DM WT protein. It could be hypothesised that the first dsRNA binding motif has an inhibitory effect on activity and requires the amino terminus to activate the domain or quench this activity. As RBM1 is capable of binding to dsRNA when expressed in isolation and did show a moderate increase in editing activity when co-expressed with the deaminase domain I would suggest that the latter is more likely.

One can only assume that the story of the dADAR dsRBD and its role has a few more chapters yet to come.

5

SUMMARY AND FUTURE WORK

5.1 SUMMARY OF WORK

Throughout this thesis I have been attempting to clarify the roles of the functional domains of the *Drosophila melanogaster* ADAR enzyme. I have demonstrated that the deaminase domain of dADAR is active and will edit both double strand RNA and *dAdar exon7* transcripts. The specificity of the deaminase domain mirrors that of the full length protein with each targeting the same adenosine bases.

The activity of the *Drosophila* ADAR protein is decreased by a serine to glycine residue change and this decrease can be replicated in a human ADAR2 mutant encoding serine at this position.

The dsRNA binding domains of dADAR share the same structural characterisitics as the domains of other well studied dsRBD-containing proteins with each being capable of binding to dsRNA on their own and when co-expressed with the deaminase domain RBM2 can increase the editing activity of the deaminase domain by around 4-fold.

5.2 FUTURE DIRECTIONS

As summarised above I have attempted to shed light on the characteristics of the *Drosophila* ADAR enzyme but as with all science, it "...never solves a

problem without creating ten more" (George Bernard Shaw) therefore the scope of possible future work is large and varied.

5.2.1 The formation of *Drosophila* ADAR dimers

The mechanism by which the *Drosophila* ADAR catalytic domain is able to edit dsRNA transcripts without the dsRNA binding domain and the dimerisation region could be investigated further. As discussed earlier, previous work suggested the formation of a homodimer was a pre-requisite for editing activity (Gallo et al., 2003). Editing activity was not detected with deletion constructs lacking the amino terminus however Gallo *et al.* assayed activity using poisoned primer extension and thin layered chromatography. In order to detect any editing activity of the deaminase domain I have to use the more sensitive method of RT-PCR. A re visit to this ΔN construct and further sequence analysis may show that it possesses very weak editing activity similar to that seen for the DM WT construct. However, the fact remains that the role of the amino terminus in dADAR is an important one and this differs from that of hADAR2 where the amino terminus appears dispensible for activity (Macbeth et al., 2004; Poulsen, 2006).

The model of dimer formation for dADAR is unclear. Poulsen *et al.* have formulated a hypothesis for the dimerisation of hADAR2 where dsRBM1 initially binds to the dsRNA and loses its auto-inhibitory effect on catalysis and allows dimers to form (Poulsen, 2006). Upon dimerisation the second dsRBM binds to the substrate resulting in a conformational change and an active enzyme. This model could be applicable to dADAR with perhaps the amino terminal playing a role in releasing a possible dsRBM1 auto inhibition in the presence of dsRNA. This hypothesis would explain the lack of editing activity by DM WT when co-expressed with the RBM1 + RBM2 with or without exon 3a as both these constructs did appear to inhibit the intrinsic deaminase domain editing activity rather than increase it.

The region of interaction between dADAR monomers has been identified but the symmetry of the dimer is unknown. Analysis of editing patterns on substrates such as the *exon 7* transcript show that more than one site is edited. Interestingly, my results show four editing sites on *exon 7* including the S/G site. If the mfold structural prediction of the secondary structure of this substrate is accurate then an adenosine opposite the S/G site is edited along with one within the tetraloop structures at either end of the duplex. This suggests a number of possibilities with regards to the quaternary structure of the homodimer but would suggest that both deaminase domains within a dimer are able to edit an adenosine.

5.2.2 The role of the edited serine residue

As previously discussed, dADAR edits its own transcript resulting in a serine to glycine change within the catalytic domain and a decrease in editing activity. This effect can be replicated in other ADARs that are not normally edited and can be mutated to code for serine at this site. So far, this auto regulatory mechanism has only been identified in *D. melanogaster* however, the idea that it may be a universal method of ADAR regulation is quashed by the fact that not all ADAR enzymes have a serine at this position; hADAR1 for example codes for aspartic acid at this position.

Considering the significant negative effect this serine to glycine change has on dADAR activity it is surprising it has not been conserved. Further investigation to identify the role of this residue would involve structural studies to ascertain if it is involved in the catalytic mechanism or perhaps interacts with the dsRNA transcript. Perhaps even more interesting would be the comparison of the hADAR1 and dADAR structures in this region to give a better understanding as to why both enzymes vary so much in what is a highly conserved region of the protein.

Appendix 1

Clustal peptide alignment of full length human ADAR2 (Accession No. NP_001103) and Drosophila ADAR (Accession No. AAF63702) proteins. All peptides referred to in the text are numbered according to the sequences below. The DM WT dADAR construct analysed is highlighted in red, the hADAR2-D construct analysed by Macbeth *et al.* is highlighted in green.

```
hADAR2          MDIEDEENMSSSSTDVKENRNLDNVSPKDGSTPGPGEGSQLSNGGGGGPGRKRPLEEGSN 60
dADAR           --MLNSANNNSPQHPVSAPSDIN-----------------MNGYNRKLPQKRGYEMPKY 40
                 .: :. * .*..  *.   :::. :....::.....:.  ** .    :**  *  .

hADAR2          GHSKYRLKKRRKTPGPVLPKNALMQLNEIKPGLQYTLLSQTGPVHAPLFVMSVEVNGQVF 120
dADAR           SDPKKKMCKERIPQ----PKNTVAMLNELRHGLIYKLESQTGPVHAPLFTISVEVDGQKY 96
                ...* :: *.* . ..  ***::  ***:: ** *.* **********.:****:** :

hADAR2          EGSGPTKKKAKLHAAEKALRSFVQFPNASEAHLAMGRTLSVNTDFTSDQADFPDTLFNGF 180
dADAR           LGQGRSKKVARIEAAATALRSFIQFKDG--AVLSP-LKPAGNLDFTSD-----EHLENGI 148
                 *.* :** *::.** .*****:** :.:.* *: . . : * *****.:. .: * **:

hADAR2          ETPDKAEPPFYVGSNGDDSFSSSGDLSLSASPVPASLAQPPLPVLPPFPPPSGKNPVMIL 240
dADAR           ENLSSSKMFEIIQTMLTEKLSNPTSLEQPTFCMSQNVSKSAITVDGQKKVPD-KGPVMLL 207
                *. ..::    : :   :.:*.. .*. .:  :. .::::.:.*      *..*.***:*

hADAR2          NELRPGLKYDFLSESGESHAKSFVMSVVVDGQFFEGSGRNKKLAKARAAQSALAAIFNLH 300
dADAR           YELFNDVNFECINIDGAQNNCRFKMTVTINEKKFDGTGPSKKTAKNAAAKAALASLCNIS 267
                 **  .:::: :. .* .:   * *:*.:: : *:*:* .** **  **::***:: *:

hADAR2          LDQTPSRQ-----PIPSEGLQLHLPQVLADAVSRLVLGKFGDLTDNFSSPHARRKVLAGV 355
dADAR           YSPMVVPQKNVPLPIDDKSSSMELPQIHADTIGRLVLEKFMEVIKGQE-AYSRRKVLAGI 326
                 .     *     ** .:. .:.***: **::.**** ** :: .. .:.::*******:

hADAR2          VMTTGTDVKDAKVISVSTGTKCINGEYMSDRGLALNDCHAEIISRRSLLRFLYTQLELYL 415
dADAR           VMTENMNFCEAKVISVSTGTKCVSGEHMSVNGAVLNDSHAEIVSRRCLLKYLYAQLDLQC 386
                *** . :. :************:.**:** .* .***.****:***.**::**:**:*

hADAR2          NNKDDQKRSIFQKSERGG--FRLKENVQFHLYISTSPCGDARIFSPHEPILEEPADRHPN 473
dADAR           N-QATAYQSIFVRNTDGQYPYKLKSGVHFHLYINTAPCGDARIFSPHEN--DTGVDKHPN 443
                *.:    :*** :.  *   ::**..*:*****.*:********.***   :  .*:***

hADAR2          RKARGQLRTKIESGEGTIPVRSNASIQTWDGVLQGERLLTMSCSDKIARWNVVGIQGSLL 533
dADAR           RKARGQLRTKIESGEGTIPVKSSDGIQTWDGVLQGQRLLTMSCSDKIARWNIVGIQGSLL 503
                ********************:*. .**********:***************:********

hADAR2          SIFVEPIYFSSIILGSLYHGDHLSRAMYQRISN-IEDLPPLYTLNKPLLSGTSNAEARQF 592
dADAR           SSIIEPVYLHSIVLGSLLHPEHMYRAVCGRIEKSIQGLPPPYHLNKPRLALVTSAEPRNQ 563
                * ::**:*: **:**** * :*: **:  **.: *:.*** * **** *: ::.**.*:

hADAR2          GKAPNFSVNWTVGDSAIEVINATTGKDELGRASRLCKHALYCRWMRVHGKVPSHLLRSKI 652
dADAR           AKAPNFGINWTIGDTELEVVNSLTGRTIGGQVSRITKQAFFVKYGFLMANLPGILVR-KV 622
                .*****.:***:**: :**:*: **:   *:.**: *:*:: ::  : .::*. *:*:*:

hADAR2          TKPNVYHESKLAAKEYQAAKARLFTAFIKAGLGAWVEKPTEQDQFSLTP 701
dADAR           TTD--YGQTKANVKDYQIAKLELFSAFKREDLGSWLKKPIEQDEFGLAE 669
                *. . * ::*  .*:** ** .**:** : .**:*::** ***:*.*:
```

BIBLIOGRAPHY

Anant, S., MacGinnitie, A.J. and Davidson, N.O. (1995) Apobec-1, the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme, is a novel RNA-binding protein. *J Biol Chem*, **270**, 14762-14767.

Arakawa, H., Hauschild, J. and Buerstedde, J.M. (2002) Requirement of the Activation-Induced Deaminase (AID) Gene for Immunoglobulin Gene Conversion. *Science*, **295**, 1301-1306.

Aruscavage, P.J. and Bass, B.L. (2000) A phylogenetic analysis reveals an unusual sequence conservation within introns involved in RNA editing. *RNA*, **6**, 257-269.

Auxilien, S., Crain, P.F., Trewyn, R.W. and Grosjean, H. (1996) Mechanism, specificity and general properties of the yeast enzyme catalysing the formation of inosine 34 in the anticodon of transfer RNA. *J. Mol. Biol.*, **262**, 437-458.

Backus, J.W. and Smith, H.C. (1991) Apolipoprotein B mRNA sequences 3' of the editing site are necessary and sufficient for editing and editosome assembly. *Nucleic Acids Res*, **19**, 6781-6786.

Backus, J.W. and Smith, H.C. (1992) Three distinct RNA sequence elements are required for efficient apolipoprotein B (apoB) RNA editing in vitro. *Nucleic Acids Res*, **20**, 6007-6014.

Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281-297.

Bass, B.L., Hurst, S.R. and Singer, J.D. (1994) Binding properties of newly identified *Xenopus* proteins containing dsRNA-binding motifs. *Curr. Biol.*, **4**, 301-314.

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

Bass, B.L., Weintraub, H., Cattaneo, R. and Billeter, M.A. (1989) Biased hypermutation of viral RNA genomes could be due to unwinding/modification of double-stranded RNA. *Cell*, **56**, 331.

Beghini, A., Ripamonti, C.B., Peterlongo, P., Roversi, G., Cairoli, R., Morra, E. and Larizza, L. (2000) RNA hyperediting and alternative splicing of

hematopoietic cell phosphatase (PTPN6) gene in acute myeloid leukemia. *Hum Mol Genet*, **9**, 2297-2304.

Benne, R., Van den Burg, J., Brakenhoff, J., 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.

Bevilacqua, P.C. and Cech, T.R. (1996) Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. *Biochemistry*, **35**, 9983-9994.

Bhalla, T., Rosenthal, J.J., Holmgren, M. and Reenan, R. (2004) Control of human potassium channel inactivation by editing of a small mRNA hairpin. *Nat Struct Mol Biol*, **11**, 950-956.

Blanc, V., Kennedy, S. and Davidson, N.O. (2003) A novel nuclear localization signal in the auxiliary domain of apobec-1 complementation factor regulates nucleocytoplasmic import and shuttling. *J Biol Chem*, **278**, 41198-41204.

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 USA*, **100**, 4102-4107.

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

Brusa, R., Zimmermann, F., Koh, D.-S., Feldmeyer, D., Gass, P., Seeburg, P.H. and Sprengel, R. (1995) Early-onset epilepsy and postnatal lethality associated with editing-deficient GluR-B allele in mice. *Science*, **270**, 1677-1680.

Burnashev, N., Monyer, H., Seeburg, P.H. and Sakmann, B. (1992) Divalent ion permeability of AMPA receptor channels is dominated by the edited form of a single subunit. *Neuron*, **8**, 189-198.

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.

Bycroft, M., Grunert, S., Murzin, A.G., Proctor, M. and St Johnston, D. (1995) NMR solution structure of a dsRNA binding domain from Drosophila staufen protein reveals homology to the N-terminal domain of ribosomal protein S5. *EMBO J*, **14**, 3563-3571.

Carter, C.W. (1998) Nucleoside deaminases for cytidine and adenosine: Comparison with deaminases acting on RNA. In Grosjean, H. and Benne, R. (eds.), *Modification and Editing of RNA.* ASM Press, Washington D.C., pp. 363-375.

Casey, J.L. and Gerin, J.L. (1995) Hepatitis D virus RNA editing: specific modification of adenosine in the antigenomic RNA. *J Virol*, **69**, 7593-7600.

Cattaneo, R., Schmid, A., Billeter, M.A., Sheppard, R.D. and Udem, S.A. (1988a) Multiple viral mutations rather than host factors cause defective measles virus gene expression in a subacute sclerosing panencephalitis cell line. *J Virol*, **62**, 1388-1397.

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

Cavaille, J., Buiting, K., Kiefmann, M., Lalande, M., Brannan, C.I., Horsthemke, B., Bachellerie, J.P., Brosius, J. and Huttenhofer, A. (2000) Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc Natl Acad Sci USA*, **97**, 14311-14316.

Chaudhuri, J., Tian, M., Khuong, C., Chua, K., Pinaud, E. and Alt, F.W. (2003) Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature*, **422**, 726-730.

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., Gotto, A.M.J., Li, W.H. and Chan, L. (1987) Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in frame stop codon. *Science*, **238**, 363-366.

Cheng, J.-F., Boyartchuk, V. and Zhu, Y. (1994) Isolation and mapping of human chromosome 21 cDNA: Progress in constructing a chromosome 21 expression map. *Genomics*, **23**, 75-84.

Cho, D.S., Yang, W., Lee, J.T., Shiekhattar, R., Murray, J.M. and Nishikura, K. (2003) Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. *J Biol Chem*, **278**, 17093-17102.

Cosentino, G.P., Venkatesan, S., Serluca, F.C., Green, S.R., Mathews, M.B. and Sonenberg, N. (1995) Double-stranded-RNA-dependent protein kinase and TAR RNA-binding protein form homo- and heterodimers in vivo. *Proc Natl Acad Sci USA*, **92**, 9445-9449.

de los Santos, T., Schweizer, J., Rees, C.A. and Francke, U. (2000) Small evolutionarily conserved RNA, resembling C/D box small nucleolar RNA, is transcribed from PWCR1, a novel imprinted gene in the Prader-Willi deletion region, which Is highly expressed in brain. *Am J Hum Genet*, **67**, 1067-1082.

Desterro, J.M., Keegan, L.P., Jaffray, E., Hay, R.T., O'Connell M, A. and Carmo-Fonseca, M. (2005) SUMO-1 modification alters ADAR1 editing activity. *Mol Biol Cell*, **16**, 5115-5126.

Desterro, J.M., Keegan, L.P., Lafarga, M., Berciano, M.T., O'Connell, M. and Carmo-Fonseca, M. (2003) Dynamic association of RNA-editing enzymes with the nucleolus. *J Cell Sci*, **116**, 1805-1818.

Dickerson, S.K., Market, E., Besmer, E. and Papavasiliou, F.N. (2003) AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med*, **197**, 1291-1296.

Dingledine, R., Borges, K., Bowie, D. and Traynelis, S.F. (1999) The glutamate receptor ion channels. *Pharmacol Rev*, **51**, 7-61.

Doi, T., Kinoshita, K., Ikegawa, M., Muramatsu, M. and Honjo, T. (2003) *De novo* protein synthesis is required for the activation-induced cytidine deaminase function in class-switch recombination. *Proc Natl Acad Sci USA*, **100**, 2634-2638.

Doyle, M. and Jantsch, M.F. (2002) New and old roles of the double-stranded RNA-binding domain. *J Struct Biol*, **140**, 147-153.

Doyle, M. and Jantsch, M.F. (2003) Distinct *in vivo* roles for double-stranded RNA-binding domains of the *Xenopus* RNA-editing enzyme ADAR1 in chromosomal targeting. *J Cell Biol*, **161**, 309-319.

Dracheva, S., Elhakem, S.L., Marcus, S.M., Siever, L.J., McGurk, S.R. and Haroutunian, V. (2003) RNA editing and alternative splicing of human serotonin 2C receptor in schizophrenia. *J Neurochem*, **87**, 1402-1412.

Driscoll, D.M., Lakhe-Reddy, S., Oleksa, L.M. and Martinez, D. (1993) Induction of RNA editing at heterologous sites by sequences in apolipoprotein B mRNA. *Mol Cell Biol*, **13**, 7288-7294.

Driscoll, D.M. and Zhang, Q. (1994) Expression and characterization of p27, the catalytic subunit of the apolipoprotein B mRNA editing enzyme. *J. Biol. Chem.*, **269**, 19843-19847.

Eckmann, C.R., Neunteufl, A., Pfaffstetter, L. and Jantsch, M.F. (2001) The human but not the *Xenopus* RNA-editing enzyme ADAR1 has an atypical nuclear localization signal and displays the characteristics of a shuttling protein. *Mol Biol Cell*, **12**, 1911-1924.

Egebjerg, J. and Heinemann, S.F. (1993) Ca2+ permeability of unedited and edited versions of the kainate selective glutamate receptor GluR6. *Proc Natl Acad Sci USA*, **90**, 755-759.

Egebjerg, J., Kukekov, V. and Heinemann, S.F. (1994) Intron sequence directs RNA editing of the glutamate receptor subunit GluR2 coding sequence. *Proc Natl Acad Sci USA*, **91**, 10270-10274.

Feldmeyer, D., Kask, K., Brusa, R., Kornau, H.C., Kolhekar, R., Rozov, A., Burnashev, N., Jensen, V., Hvalby, O., Sprengel, R. and Seeburg, P.H. (1999) Neurological dysfunctions in mice expressing different levels of the Q/R site-unedited AMPAR subunit GluR-B. *Nat Neurosci*, **2**, 57-64.

Fitzgerald, L.W., Iyer, G., Conklin, D.S., Krause, C.M., Marshall, A., Patterson, J.P., Tran, D.P., Jonak, G.J. and Hartig, P.R. (1999) Messenger RNA Editing of the Human Serotonin 5-HT(2C) Receptor. *Neuropsychopharmacology*, **21 Suppl 1**, S82-90.

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.

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.

George, C.X., Wagner, M.V. and Samuel, C.E. (2005) Expression of interferon-inducible RNA adenosine deaminase ADAR1 during pathogen infection and mouse embryo development involves tissue-selective promoter utilization and alternative splicing. *J Biol Chem*, **280**, 15020-15028.

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., O'Connell, M.A. and Keller, W. (1997) Two forms of human double-stranded RNA-specific editase 1 (hRED1) generated by the insertion of an Alu cassette. *RNA*, **3**, 453-463.

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.

Gohla, A., Offermanns, S., Wilkie, T.M. and Schultz, G. (1999) Differential involvement of Galpha12 and Galpha13 in receptor-mediated stress fiber formation. *J Biol Chem*, **274**, 17901-17907.

Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B. (2002) Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing. *Genetics*, **160**, 1519-1533.

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.

Grosjean, H. and Benne, R. (1998) *Modification and Editing of RNA*. ASM Press, Washington, DC.

Gurevich, I., Englander, M.T., Adlersberg, M., Siegal, N.B. and Schmauss, C. (2002a) Modulation of serotonin 2C receptor editing by sustained changes in serotonergic neurotransmission. *J Neurosci*, **22**, 10529-10532.

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

Hajjar, A.M. and Linial, M.L. (1995) Modification of retroviral RNA by double-stranded RNA adenosine deaminase. *J Virol*, **69**, 5878-5882.

Hall, W.W., Lamb, R.A. and Choppin, P.W. (1979) Measles and subacute sclerosing panencephalitis virus proteins: lack of antibodies to the M

protein in patients with subacute sclerosing panencephalitis. *Proc Natl Acad Sci USA*, **76**, 2047-2051.

Hanrahan, C.J., Palladino, M.J., Bonneau, L.J. and Reenan, R.A. (1998) RNA Editing of a Drosophila Sodium Channel Gene. *Annals of the New York Academy of Science*, **868**, 51-66.

Hanrahan, C.J., Palladino, M.J., Ganetzky, B. and Reenan, R.A. (2000) RNA editing of the drosophila para Na(+) channel transcript. Evolutionary conservation and developmental regulation. *Genetics*, **155**, 1149-1160.

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.

Herb, A., Higuchi, M., Sprengel, R. and Seeburg, P.H. (1996) Q/R site editing in kainate receptor GluR5 and GluR6 pre-mRNAs requires distant intronic sequences. *Proc Natl Acad Sci USA*, **93**, 1875-1880.

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 USA*, **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 USA*, **98**, 12132-12137.

Herbert, A., Schade, M., Lowenhaupt, K., Alfken, J., Schwartz, T., Shlyakhtenko, L.S., Lyubchenko, Y.L. and Rich, A. (1998) The Zalpha domain from human ADAR1 binds to the Z-DNA conformer of many different sequences. *Nucleic Acids Res*, **26**, 3486-3493.

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., Köhler, 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.

Hirano, K., Young, S.G., Farese, R.V., Jr., Ng, J., Sande, E., Warburton, C., Powell-Braxton, L.M. and Davidson, N.O. (1996) Targeted disruption

of the mouse apobec-1 gene abolishes apolipoprotein B mRNA editing and eliminates apolipoprotein B48. *J Biol Chem*, **271**, 9887-9890.

Hitti, E.G., Sallacz, N.B., Schoft, V.K. and Jantsch, M.F. (2004) Oligomerization activity of a double-stranded RNA-binding domain. *FEBS Lett*, **574**, 25-30.

Hooks, J.J., Moutsopoulos, H.M., Geis, S.A., Stahl, N.I., Decker, J.L. and Notkins, A.L. (1979) Immune interferon in the circulation of patients with autoimmune disease. *N Engl J Med*, **301**, 5-8.

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

Horikami, S.M. and Moyer, S.A. (1995) Double-stranded RNA adenosine deaminase activity during measles virus infection. *Virus Res*, **36**, 87-96.

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.

Hurst, S.R., Hough, R.F., Aruscavage, P.J. and Bass, B.L. (1995) Deamination of mammalian glutamate receptor RNA by *Xenopus* dsRNA adenosine deaminase: Similarties to *in vivo* RNA editing. *RNA*, **1**, 1051-1060.

Innerarity, T.L., Borén, J., Yamanaka, S. and Olofsson, S.-O. (1996) Biosynthesis of apolipoprotein B48-containing lipoproteins. *J Biol Chem*, **271**, 2353-2356.

Ito, S., Nagaoka, H., Shinkura, R., Begum, N., Muramatsu, M., Nakata, M. and Honjo, T. (2004) Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc Natl Acad Sci USA*, **101**, 1975-1980.

Iwamoto, K. and Kato, T. (2003) RNA editing of serotonin 2C receptor in human postmortem brains of major mental disorders. *Neurosci Lett*, **346**, 169-172.

Iwamoto, K., Nakatani, N., Bundo, M., Yoshikawa, T. and Kato, T. (2005) Altered RNA editing of serotonin 2C receptor in a rat model of depression. *Neurosci Res*, **53**, 69-76.

Jaikaran, D.C., Collins, C.H. and MacMillan, A.M. (2002) Adenosine to inosine editing by ADAR2 requires formation of a ternary complex on the GluR-B R/G site. *J Biol Chem*, **277**, 37624-37629.

Jayan, G.C. and Casey, J.L. (2002a) Increased RNA editing and inhibition of hepatitis delta virus replication by high-level expression of ADAR1 and ADAR2. *J Virol*, **76**, 3819-3827.

Jayan, G.C. and Casey, J.L. (2002b) Inhibition of hepatitis delta virus RNA editing by short inhibitory RNA-mediated knockdown of ADAR1 but not ADAR2 expression. *J Virol*, **76**, 12399-12404.

Jonas, P. and Sakmann, B. (1992) Glutamate receptor channels in isolated patches from CA1 and CA3 pyramidal cells of rat hippocampal slices. *J Physiol*, **455**, 143-171.

Kask, K., Zamanillo, D., Rozov, A., Burnashev, N., Sprengel, R. and Seeburg, P.H. (1998) The AMPA receptor subunit GluR-B in its Q/R site-unedited form is not essential for brain development and function. *Proc Natl Acad Sci USA*, **95**, 13777-13782.

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.

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., 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.

Kharrat, A., Macias, M.J., Gibson, T.J., Nilges, M. and Pastore, A. (1995) Structure of the dsRNA binding domain of *E. coli* RNase III. *EMBO J*, **14**, 3572-3584.

Kim, D.D., Kim, T.T., Walsh, T., Kobayashi, Y., Matise, T.C., Buyske, S. and Gabriel, A. (2004) Widespread RNA editing of embedded alu elements in the human transcriptome. *Genome Res*, **14**, 1719-1725.

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 cDNAs for double-stranded RNA adenosine deaminase, a

candidate enzyme for nuclear RNA editing. *Proc Natl Acad Sci USA*, **91**, 11457-11461.

Kimelman, D. and Kirschner, M.W. (1989) An antisense mRNA directs the covalent modification of the transcript encoding fibroblast growth factor in *Xenopus* oocytes. *Cell*, **59**, 687-696.

Kiss, T. (2001) NEW EMBO MEMBER'S REVIEW: Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J*, **20**, 3617-3622.

Knight, S.W. and Bass, B.L. (2002) The role of RNA editing by ADARs in RNAi. *Mol Cell*, **10**, 809-817.

Köhler, M., Burnashev, N., Sakmann, B. and Seeburg, P.H. (1993) Determinants of Ca^{2+} permeability in both TM1 and TM2 of high affinity kainate receptor channels: Diversity by RNA editing. *Neuron*, **10**, 491-500.

Krovat, B.C. and Jantsch, M.F. (1996) Comparative mutational analysis of the double-stranded RNA binding domains of *Xenopus laevis* RNA-binding protein A. *J Biol Chem*, **271**, 28112-28119.

Kumar, M. and Carmichael, G.C. (1997) Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *Proc Natl Acad Sci USA*, **94**, 3542-3547.

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., Chen, C.-X., Carter, K.C. and Nishikura, K. (1997) Editing of glutamate receptor B subunit ion channel RNAs by four alternatively spliced DRADA2 double-stranded RNA adenosine deaminases. *Mol Cell Biol*, **17**, 2413-2424.

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.

Lai, M.M. (1995) The molecular biology of hepatitis delta virus. *Annu Rev Biochem*, **64**, 259-286.

Lamontagne, B., Tremblay, A. and Abou Elela, S. (2000) The N-terminal domain that distinguishes yeast from bacterial RNase III contains a dimerization signal required for efficient double-stranded RNA cleavage. *Mol Cell Biol*, **20**, 1104-1115.

Lau, P.P., Zhu, H.J., Baldini, A., Charnsangavej, C. and Chan, L. (1994) Dimeric structure of a human apolipoprotein B mRNA editing protein and cloning and chromosomal localization of its gene. *Proc Natl Acad Sci USA*, **91**, 8522-8526.

Laxminarayana, D., Khan, I.U. and Kammer, G. (2002) Transcript mutations of the alpha regulatory subunit of protein kinase A and up-regulation of the RNA-editing gene transcript in lupus T lymphocytes. *Lancet*, **360**, 842-849.

Lee, C.J., Kong, H., Manzini, M.C., Albuquerque, C., Chao, M.V. and MacDermott, A.B. (2001) Kainate receptors expressed by a subpopulation of developing nociceptors rapidly switch from high to low Ca2+ permeability. *J Neurosci*, **21**, 4572-4581.

Lehmann, K.A. and Bass, B.L. (1999) The importance of internal loops within RNA substrates of ADAR1. *J Mol Biol*, **291**, 1-13.

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

Lellek, H., Kirsten, R., Diehl, I., Apostel, F., Buck, F. and Greeve, J. (2000) Purification and molecular cloning of a novel essential component of the apolipoprotein B mRNA editing enzyme-complex. *J Biol Chem*, **275**, 19848-19856.

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.

Liu, Y., George, C.X., Patterson, J.B. and Samuel, C. (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 USA*, **97**, 12541-12546.

Liu, Y. and Samuel, C.E. (1996) Mechanism of interferon action: Functionally distinct RNA-binding and catalytic domains in the interferon-

inducible, double-stranded RNA-specific adenosine deaminase. *J Virol*, **70**, 1961-1968.

Liu, Y., Wolff, K.C., Jacobs, B.L. and Samuel, C.E. (2001) Vaccinia virus E3L interferon resistance protein inhibits the interferon-induced adenosine deaminase A-to-I editing activity. *Virology*, **289**, 378-387.

Lomeli, H., Mosbacher, J., Melcher, T., Höger, 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.

Loughney, K., Kreber, R. and Ganetzky, B. (1989) Molecular analysis of the para locus, a sodium channel gene in *Drosophila*. *Cell*, **58**, 1143-1154.

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

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.

Macbeth, M.R., Lingam, A.T. and Bass, B.L. (2004) Evidence for auto-inhibition by the N terminus of hADAR2 and activation by dsRNA binding. *Rna*, **10**, 1563-1571.

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.

MacGinnitie, A.J., Anant, S. and Davidson, N.O. (1995) Mutagenesis of apobec-1, the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme, reveals distinct domains that mediate cytosine nucleoside deaminase, RNA binding, and RNA editing activity. *J Biol Chem*, **270**, 14768-14775.

Martinez, I., Dopazo, J. and Melero, J.A. (1997) Antigenic structure of the human respiratory syncytial virus G glycoprotein and relevance of hypermutation events for the generation of antigenic variants. *J Gen Virol*, **78 (Pt 10)**, 2419-2429.

Martinez, I. and Melero, J.A. (2002) A model for the generation of multiple A to G transitions in the human respiratory syncytial virus genome: predicted RNA secondary structures as substrates for adenosine deaminases that act on RNA. *J Gen Virol*, **83**, 1445-1455.

McGrew, L., Price, R.D., Hackler, E., Chang, M.S. and Sanders-Bush, E. (2004) RNA Editing of the Human Serotonin 5-HT2C Receptor Disrupts Transactivation of the Small G-Protein RhoA. *Mol Pharmacol*, **65**, 252-256.

McMillan, N.A., Carpick, B.W., Hollis, B., Toone, W.M., Zamanian-Daryoush, M. and Williams, B.R. (1995) Mutational analysis of the double-stranded RNA (dsRNA) binding domain of the dsRNA-activated protein kinase, PKR. *J Biol Chem*, **270**, 2601-2606.

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., Higuchi, M. and Seeburg, P.H. (1996a) RED2, a brain specific member of the RNA-specific adenosine deaminase family. *J Biol Chem*, **271**, 31795-31798.

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

Melcher, T., Maas, S., Higuchi, M., Keller, W. and Seeburg, P.H. (1995) Editing of AMPA receptor GluR-B pre-mRNA in vitro reveals site-selective adenosine to inosine conversion. *J Biol Chem*, **270**, 8566-8570.

Micklem, D.R., Adams, J., Grunert, S. and St Johnston, D. (2000) Distinct roles of two conserved Staufen domains in oskar mRNA localization and translation. *Embo J*, **19**, 1366-1377.

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.

Morrison, J.R., Paszty, C., Stevens, M.E., Hughes, S.D., Forte, T., Scott, J. and Rubin, E.M. (1996) Apolipoprotein B RNA editing enzyme-deficient

mice are viable despite alterations in lipoprotein metabolism. *Proc Natl Acad Sci USA*, **93**, 7154-7159.

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 USA*, **99**, 7906-7911.

Morse, D.P. and Bass, B.L. (1999) Long RNA hairpins that contain inosine are present in Caenorhabditis elegans poly(A)+ RNA. *Proc Natl Acad Sci USA*, **96**, 6048-6053.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y. and Honjo, T. (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, **102**, 553-563.

Muramatsu, M., Sankaranand, V.S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O. and Honjo, T. (1999) Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem*, **274**, 18470-18476.

Muto, T., Muramatsu, M., Taniwaki, M., Kinoshita, K. and Honjo, T. (2000) Isolation, tissue distribution, and chromosomal localization of the human activation-induced cytidine deaminase (AID) gene. *Genomics*, **68**, 85-88.

Nagaoka, H., Muramatsu, M., Yamamura, N., Kinoshita, K. and Honjo, T. (2002) Activation-induced deaminase (AID)-directed hypermutation in the immunoglobulin Smu region: implication of AID involvement in a common step of class switch recombination and somatic hypermutation. *J Exp Med*, **195**, 529-534.

Nakamuta, M., Chang, B.H., Zsigmond, E., Kobayashi, K., Lei, H., Ishida, B.Y., Oka, K., Li, E. and Chan, L. (1996) Complete phenotypic characterization of apobec-1 knockout mice with a wild-type genetic background and a human apolipoprotein B transgenic background, and restoration of apolipoprotein B mRNA editing by somatic gene transfer of Apobec-1. *J Biol Chem*, **271**, 25981-25988.

Nanduri, S., Carpick, B.W., Yang, Y., Williams, B.R. and Qin, J. (1998) Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. *EMBO J*, **17**, 5458-5465.

Navaratnam, N., Bhattacharya, S., Fujino, T., Patel, D., Jarmuz, A.L. and Scott, J. (1995) Evolutionary origins of apoB mRNA editing: catalysis by a cytidine deaminase that has acquired a novel RNA-binding motif at its active site. *Cell*, **81**, 187-195.

Navaratnam, N., Morrison, J.R., Bhattacharya, S., Patel, D., Funahashi, T., Giannoni, F., Teng, B., Davidson, N.O. and Scott, J. (1993) The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *J Biol Chem*, **268**, 20709-20712.

Nie, Y., Zhao, Q., Su, Y. and Yang, J.H. (2004) Subcellular distribution of ADAR1 isoforms is synergistically determined by three nuclear discrimination signals and a regulatory motif. *J Biol Chem*.

Nishikura, K., Yoo, C., Kim, U., Murray, J.M., Estes, P.A., Cash, F.E. and Liebhaber, S.A. (1991) Substrate specificity of the dsRNA unwinding/modifying activity. *EMBO J.*, **10**, 3523-3532.

Niswender, C.M., Copeland, S.C., Herrick-Davis, K., Emeson, R.B. and Sanders-Bush, E. (1999) RNA editing of the human serotonin 5-hydroxytryptamine 2C receptor silences constitutive activity. *J Biol Chem*, **274**, 9472-9478.

Niswender, C.M., Herrick-Davis, K., Dilley, G.E., Meltzer, H.Y., Overholser, J.C., Stockmeier, C.A., Emeson, R.B. and Sanders-Bush, E. (2001) RNA editing of the human serotonin 5-HT2C receptor. Alterations in suicide and implications for serotonergic pharmacotherapy. *Neuropsychopharmacology*, **24**, 478-491.

O'Connell, M.A., Gerber, A. and Keller, W. (1997) Purification of human double-stranded RNA-specific editase1 (hRed1), involved in editing of brain glutamate receptor B pre-mRNA. *J Biol Chem*, **272**, 473-478.

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 USA*, **91**, 10596-10600.

O'Connell, M.A., Krause, S., Higuchi, M., Hsuan, J.J., Totty, N.F., Jenny, A. and Keller, W. (1995) Cloning of cDNAs encoding mammalian double-stranded RNA-specific adenosine deaminase. *Mol Cell Biol*, **15**, 1389-1397.

O'Hara, P.J., Nichol, S.T., Horodyski, F.M. and Holland, J.J. (1984) Vesicular stomatitis virus defective interfering particles can contain extensive genomic sequence rearrangements and base substitutions. *Cell*, **36**, 915-924.

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.

Patel, R.C. and Sen, G.C. (1998) Requirement of PKR dimerization mediated by specific hydrophobic residues for its activation by double-stranded RNA and its antigrowth effects in yeast. *Mol Cell Biol*, **18**, 7009-7019.

Patel, R.C., Stanton, P., McMillan, N.M., Williams, B.R. and Sen, G.C. (1995) The interferon-inducible double-stranded RNA-activated protein kinase self-associates *in vitro* and *in vivo*. *Proc Natl Acad Sci USA*, **92**, 8283-8287.

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

Patterson, J.B., Thomis, D.C., Hans, S.L. and Samuel, C.E. (1995) Mechanism of interferon action: double-stranded RNA-specific adenosine deaminase from human cells is inducible by alpha and gamma interferons. *Virology*, **210**, 508-511.

Peixoto, A.A. and Hall, J.C. (1998) Analysis of temperature-sensitive mutants reveals new genes involved in the courtship song of *Drosophila*. *Genetics*, **148**, 827-838.

Peixoto, A.A., Smith, L.A. and Hall, J.C. (1997) Genomic organization and evolution of alternative exons in a *Drosophila* calcium channel gene. *Genetics*, **145**, 1003-1013.

Peters, N.T., Rohrbach, J.A., Zalewski, B.A., Byrkett, C.M. and Vaughn, J.C. (2003) RNA editing and regulation of *Drosophila 4f-rnp* expression by *sas-10* antisense readthrough mRNA transcripts. *RNA*, **9**, 698-710.

Petersen-Mahrt, S.K., Harris, R.S. and Neuberger, M.S. (2002) AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature*, **418**, 99-103.

Petschek, J.P., Mermer, M.J., Scheckelhoff, M.R., Simone, A.A. and Vaughn, J.C. (1996) RNA editing in *Drosophila 4f-rnp* gene nuclear transcripts by multiple A-to-G conversions. *J Mol Biol*, **259**, 885-890.

Petschek, J.P., Scheckelhoff, M.R., Mermer, M.J. and Vaughn, J.C. (1997) RNA editing and alternative splicing generate mRNA transcript diversity from the *Drosophila 4f-rnp* locus. *Gene*, **204**, 267-276.

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

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.

Polson, A.G., Ley, H.L., 3rd, Bass, B.L. and Casey, J.L. (1998) Hepatitis delta virus RNA editing is highly specific for the amber/W site and is suppressed by hepatitis delta antigen. *Mol Cell Biol*, **18**, 1919-1926.

Poulsen, H., Jorgensen, R., Heding, A., Nielsen, F.C., Bonven, B., and Egebjerg, J. (2006) Dimerization of ADAR2 is mediated by the double-stranded RNA binding domain. *RNA*, **In press**.

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-48 in intestine. *Cell*, **50**, 831-840.

Price, R.D., Weiner, D.M., Chang, M.S. and Sanders-Bush, E. (2001) RNA editing of the human serotonin 5-HT2C receptor alters receptor-mediated activation of G13 protein. *J Biol Chem*, **276**, 44663-44668.

Rabinovici, R., Kabir, K., Chen, M., Su, Y., Zhang, D., Luo, X. and Yang, J.H. (2001) ADAR1 is involved in the development of microvascular lung injury. *Circ Res*, **88**, 1066-1071.

Raitskin, O., Cho, D.S., Sperling, J., Nishikura, K. and Sperling, R. (2001) RNA editing activity is associated with splicing factors in lnRNP particles: The nuclear pre-mRNA processing machinery. *Proc Natl Acad Sci USA*, **98**, 6571-6576.

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.

Reenan, R.A. (2005) Molecular determinants and guided evolution of species-specific RNA editing. *Nature*, **434**, 409-413.

Reenan, R.A., Hanrahan, C.J. and Barry, G. (2000) The mle(napts) RNA helicase mutation in *Drosophila* results in a splicing catastrophe of the *para* Na+ channel transcript in a region of RNA editing. *Neuron*, **25**, 139-149.

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Labelouse, R., Gennery, A., Tezcan, I., Ersoy, F., Kayserili, H., Ugazio, A.G., Brousse, N., Muramatsu, M., Notarangelo, L.D., Kinoshita, K., Honjo, T., Fischer, A. and Durandy, A. (2000) Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell*, **102**, 565-575.

Romano, P.R., Zhang, F., Tan, S.L., Garcia-Barrio, M.T., Katze, M.G., Dever, T.E. and Hinnebusch, A.G. (1998) Inhibition of double-stranded RNA-dependent protein kinase PKR by vaccinia virus E3: role of complex formation and the E3 N-terminal domain. *Mol Cell Biol*, **18**, 7304-7316.

Ruano, D., Lambolez, B., Rossier, J., Paternain, A.V. and Lerma, J. (1995) Kainate receptor subunits expressed in single cultured hippocampal neurons: molecular and functional variants by RNA editing. *Neuron*, **14**, 1009-1017.

Rueter, S.M., Burns, C.M., Coode, S.A., Mookherjee, P. and Emeson, R.B. (1995) Glutamate receptor RNA editing *in vitro* by enzymatic conversion of adenosine to inosine. *Science*, **267**, 1491-1494.

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

Ryter, J.M. and Schultz, S.C. (1998) Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.*, **17**, 7505-7513.

Saccomanno, L. and Bass, B.L. (1999) A minor fraction of basic fibroblast growth factor mRNA is deaminated in Xenopus stage VI and matured oocytes. *RNA*, **5**, 39-48.

Sansam, C.L., Wells, K.S. and Emeson, R.B. (2003) Modulation of RNA editing by functional nucleolar sequestration of ADAR2. *Proc Natl Acad Sci U S A*.

Sato, S., Wong, S.K. and Lazinski, D.W. (2001) Hepatitis delta virus minimal substrates competent for editing by ADAR1 and ADAR2. *J Virol*, **75**, 8547-8555.

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 O'Connell, M.A. (2005) Cleavage of dsRNAs hyper-edited by ADARs occurs at preferred editing sites. *Nucleic Acids Res*, **33**, 5954-5964.

Scadden, A.D. and Smith, C.W. (1997) A ribonuclease specific for inosine-containing RNA: a potential role in antiviral defence? *EMBO J*, **16**, 2140-2149.

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

Schmauss, C. (2003) Serotonin 2C receptors: suicide, serotonin, and runaway RNA editing. *Neuroscientist*, **9**, 237-242.

Schock, D., Kuo, S.R., Steinburg, M.F., Bolognino, M., Sparks, J.D., Sparks, C.E. and Smith, H.C. (1996) An auxiliary factor containing a 240-kDa protein complex is involved in apolipoprotein B RNA editing. *Proc. Natl. Acad. Sci. USA*, **93**, 1097-1102.

Schwartz, T., Lowenhaupt, K., Kim, Y.G., Li, L., Brown, B.A., 2nd, Herbert, A. and Rich, A. (1999) Proteolytic dissection of Zab, the Z-DNA-binding domain of human ADAR1. *J Biol Chem*, **274**, 2899-2906.

Semenov, E.P. and Pak, W.L. (1999) Diversification of *Drosophila* chloride channel gene by multiple posttranscriptional mRNA modifications. *J Neurochem*, **72**, 66-72.

Shah, R.R., Knott, T.J., Legros, J.E., Navaratnam, N., Greeve, J.C. and Scott, J. (1991) Sequence requirements for the editing of apolipoprotein B mRNA. *J Biol Chem*, **266**, 16301-16304.

Smith, L.A., Peixoto, A.A. and Hall, J.C. (1998) RNA editing in the *Drosophila DMCA1A* calcium-channel alpha 1 subunit transcript. *J Neurogenet*, **12**, 227-240.

Smith, L.A., Wang, X.J., Peixoto, A.A., Neumann, E.K., Hall, L.M. and Hall, J.C. (1996) A *Drosophila* calcium channel α1 subunit gene maps to a genetic locus associated with behavioural and visual defects. *J. Neuroscience*, **16**, 7868-7879.

Sodhi, M.S., Burnet, P.W., Makoff, A.J., Kerwin, R.W. and Harrison, P.J. (2001) RNA editing of the 5-HT(2C) receptor is reduced in schizophrenia. *Mol Psychiatry*, **6**, 373-379.

Sommer, B., Köhler, 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.

Spanggord, R.J. and Beal, P.A. (2001) Selective binding by the RNA binding domain of PKR revealed by affinity cleavage. *Biochemistry*, **40**, 4272-4280.

St Johnston, D., Brown, N.H., Gall, J.G. and Jantsch, M. (1992) A conserved double-stranded RNA-binding domain. *Proc Natl Acad Sci USA*, **89**, 10979-10983.

Stefl, R. and Allain, F.H. (2005) A novel RNA pentaloop fold involved in targeting ADAR2. *RNA*, **11**, 592-597.

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.

Steitz, T.A. (1993) Similarities and differences between RNA and DNA recognition by proteins. In Gesteland, R.F. and Atkins, J.F. (eds.), *"The RNA World"*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., pp. 219-237.

Stephens, O.M., Haudenschild, B.L. and Beal, P.A. (2004) The Binding Selectivity of ADAR2's dsRBMs Contributes to RNA-Editing Selectivity. *Chem Biol*, **11**, 1239-1250.

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.

Tanoue, A., Koshimizu, T.A., Tsuchiya, M., Ishii, K., Osawa, M., Saeki, M. and Tsujimoto, G. (2002) Two novel transcripts for human endothelin B receptor produced by RNA editing /alternative splicing from a single gene. *J Biol Chem*, **21**, 21.

Teng, B., Burant, C.F. and Davidson, N.O. (1993) Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science*, **260**, 1816-1819.

Teng, B., Ishida, B., Forte, T.M., Blumenthal, S., Song, L.Z., Gotto, A.M., Jr. and Chan, L. (1997) Effective lowering of plasma, LDL, and esterified cholesterol in LDL receptor-knockout mice by adenovirus-mediated

gene delivery of ApoB mRNA editing enzyme (Apobec1). *Arterioscler Thromb Vasc Biol*, **17**, 889-897.

Tian, B. and Mathews, M.B. (2001) Functional characterization of and cooperation between the double-stranded RNA-binding motifs of the protein kinase PKR. *J Biol Chem*, **276**, 9936-9944.

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.

Ung, T.L., Cao, C., Lu, J., Ozato, K. and Dever, T.E. (2001) Heterologous dimerization domains functionally substitute for the double-stranded RNA binding domains of the kinase PKR. *EMBO J*, **20**, 3728-3737.

Vissel, B., Royle, G.A., Christie, B.R., Schiffer, H.H., Ghetti, A., Tritto, T., Perez-Otano, I., Radcliffe, R.A., Seamans, J., Sejnowski, T., Wehner, J.M., Collins, A.C., O'Gorman, S. and Heinemann, S.F. (2001) The role of RNA editing of kainate receptors in synaptic plasticity and seizures. *Neuron*, **29**, 217-227.

Vitali, P., Basyuk, E., Le Meur, E., Bertrand, E., Muscatelli, F., Cavaille, J. and Huttenhofer, A. (2005) ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs. *J Cell Biol*, **169**, 745-753.

Wagner, R.W., Yoo, C., Wrabetz, L., Kamholz, J., Buchhalter, J., Hassan, N.F., Khalili, K., Kim, S.U., Perussia, B., McMorris, F.A. and Nishikura, K. (1990) Double-stranded RNA unwinding and modifying activity is detected ubiquitously in primary tissues and cell lines. *Mol Cell Biol*, **10**, 5586-5590.

Wang, Q., Khillan, J., Gadue, P. and Nishikura, K. (2000) Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. *Science*, **290**, 1765-1768.

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

Wang, Y., Zeng, Y., Murray, J.M. and Nishikura, K. (1995) Genomic organization and chromosomal location of the human dsRNA

adenosine deaminase gene: the enzyme for glutamate-activated ion channel RNA editing. *J Mol Biol*, **254**, 184-195.

Wolf, J., Gerber, A.P. and Keller, W. (2002) tadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*. *EMBO J*, **21**, 3841-3851.

Wong, S.K. and Lazinski, D.W. (2002) Replicating hepatitis delta virus RNA is edited in the nucleus by the small form of ADAR1. *Proc Natl Acad Sci USA*, **99**, 15118-15123.

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

Wong, T.C., Ayata, M., Hirano, A., Yoshikawa, Y., Tsuruoka, H. and Yamanouchi, K. (1989) Generalized and localized biased hypermutation affecting the matrix gene of a measles virus strain that causes subacute sclerosing panencephalitis. *J Virol*, **63**, 5464-5468.

Wong, T.C., Ayata, M., Ueda, S. and Hirano, A. (1991) Role of biased hypermutation in evolution of subacute sclerosing panencephalitis virus from progenitor acute measles virus. *J Virol*, **65**, 2191-2199.

Wu, H., Henras, A., Chanfreau, G. and Feigon, J. (2004) Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of *Rnt1p* RNase III. *Proc Natl Acad Sci USA*, **101**, 8307-8312.

Wu, S. and Kaufman, R.J. (1997) A model for the double-stranded RNA (dsRNA)-dependent dimerization and activation of the dsRNA-activated protein kinase PKR. *J Biol Chem*, **272**, 1291-1296.

Yamanaka, S., Balestra, M.E., Ferrell, L.D., Fan, J., Arnold, K.S., Taylor, S., Taylor, J.M. and Innerarity, T.L. (1995) Apolipoprotein B mRNA-editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals. *Proc Natl Acad Sci USA*, **92**, 8483-8487.

Yamanaka, S., Poksay, K.S., Arnold, K.S. and Innerarity, T.L. (1997) A novel translational repressor mRNA is edited extensively in livers containing tumors caused by the transgene expression of the apoB mRNA-editing enzyme. *Genes Dev*, **11**, 321-333.

Yang, J.H., Luo, X., Nie, Y., Su, Y., Zhao, Q., Kabir, K., Zhang, D. and Rabinovici, R. (2003) Widespread inosine-containing mRNA in lymphocytes regulated by ADAR1 in response to inflammation. *Immunology*, **109**, 15-23.

Yang, W., Wang, Q., Kanes, S.J., Murray, J.M. and Nishikura, K. (2004) Altered RNA editing of serotonin 5-HT(2C) receptor induced by

interferon: implications for depression associated with cytokine therapy. *Brain Res Mol Brain Res*, **124**, 70-78.

Yi-Brunozzi, H.Y., Easterwood, L.M., Kamilar, G.M. and Beal, P.A. (1999) Synthetic substrate analogs for the RNA-editing adenosine deaminase ADAR-2. *Nucleic Acids Res*, **27**, 2912-2917.

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.

Zhang, F., Romano, P.R., Nagamura-Inoue, T., Tian, B., Dever, T.E., Mathews, M.B., Ozato, K. and Hinnebusch, A.G. (2001) Binding of double-stranded RNA to protein kinase PKR is required for dimerization and promotes critical autophosphorylation events in the activation loop. *J Biol Chem*, **276**, 24946-24958.

Zhang, X.J., He, P.P., Li, M., He, C.D., Yan, K.L., Cui, Y., Yang, S., Zhang, K.Y., Gao, M., Chen, J.J., Li, C.R., Jin, L., Chen, H.D., Xu, S.J. and Huang, W. (2004) Seven novel mutations of the ADAR gene in Chinese families and sporadic patients with dyschromatosis symmetrica hereditaria (DSH). *Hum Mutat*, **23**, 629-630.

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.

Zhu, Y., Nonoyama, S., Morio, T., Muramatsu, M., Honjo, T. and Mizutani, S. (2003) Type two hyper-IgM syndrome caused by mutation in activation-induced cytidine deaminase. *J Med Dent Sci*, **50**, 41-46.

Zuker, M. (1989) Computer prediction of RNA structure. *Meth Enzymol*, **180**, 262-288.