7 GENERAL DISCUSSION

In this thesis, a survey of the types and targets of RNA editing in the human brain is presented. Approximately 1 in 1,700 nucleotides in the human brain RNA sample used were subject to A > I editing. By contrast, RNA editing by mechanisms other than A > I is a rare event in the human brain. The majority of A > I edits are in transcribed intronic and intergenic Alu repeats, and are associated with dsRNA formation with inverted Alus in the same transcript. Within edited Alu sequences, A > I editing occurs preferentially at adenosines with a deficit of guanine at the immediately 5' adjacent nucleotide, and an increase in guanine at the immediately 3' adjacent nucleotide. Editing is also more efficient at A:C mismatches than at other mismatches or A:U matches in simulations of dsRNA. The results suggest that the effect of A > I editing is to increase the number of mismatches in dsRNA molecules, albeit by a relatively modest amount (in edited sequences, an additional 1-2% of base pairs become mismatched after editing).

7.1 FUTURE CHALLENGES

We cannot currently rule out the existence of non A > I RNA edits in human brain RNA. The scarcity of such edits means that evaluation of additional sequence variants from a more extensive survey of the type described in this thesis or by a targeted approach such as RT-PCR product sequencing will be necessary for their identification. A more extensive survey would allow the frequency with which such RNA edits occur in the human brain to be determined more accurately. It is also possible that human brain transcripts harbour additional coding RNA edits. A more exhaustive investigation directed

at coding sequences is warranted to detect rare, functionally important coding edits. This could be achieved by sequencing from cDNA clones derived from cytoplasmic RNA, or by sequencing RT-PCR products designed to amplify specifically from coding sequences. Experimental analysis of the exonic Alu sequences with an inverted copy in an adjacent intron identified in this survey may also reveal novel coding A > I edits.

This survey was performed on poly (A)+ RNA. Further work is required to investigate the extent to which non-coding unadenylated RNAs are subject to RNA editing. The function of many non-coding RNAs is dependent on base pairing and local dsRNA structures, and may plausibly be regulated by RNA editing. The presence of known A > I edits in tRNAs (Maas et al., 1999) and miRNAs (Luciano et al., 2004) are further indications that a survey of non-coding RNAs is warranted.

Several analyses indicate that A > I editing varies widely between different tissues (Paul and Bass, 1998, Levanon et al., 2004, Kim et al., 2004). It will be interesting to carry out a more exhaustive analysis of the patterns of A > I editing in different tissues, and to look for correlation with the expression levels of the different ADAR editing enzymes in these tissues. More extensive evaluation of the patterns of A > I editing, and ADAR expression in diseased tissues is also warranted, as aberrant A > I editing has previously been linked with tumour progression in gliomas (Maas et al., 2001b), and a number of neurological disorders including amyotrophic lateral sclerosis (ALS) and epilepsy (Kawahara et al., 2004, Kortenbruck et al., 2001). As C > U RNA

editing of ApoB mRNA occurs specifically expressed in the small intestine (Teng et al., 1993), it is possible that additional tissue specific RNA editing activities may exist. This could be assessed by performing a survey of RNA editing in other tissues similar to the one described in this thesis.

In our analyses of A > I editing from total cDNA, we found that the extent of A > I editing varied between different adenosines in the same transcript. This suggests that within the total population of transcripts, individual molecules are differently edited. Cloning and sequencing of multiple individual cDNAs from the same transcript will be required to better understand the patterns of A > I editing at the level of individual RNA molecules.

Currently, the extent to which each of the ADAR editing enzymes contributes to the pattern of A > I edits observed in this survey is unknown. One way of investigating this further would be to use RNA interference to selectively down-regulate ADAR1 or ADAR2 in cultured cells in order to investigate the contribution of each enzyme to the pattern of A > I edits identified in Alu sequences from this survey. This type of analysis may also help elucidate the functional consequences of Alu A > I editing.

7.2 THE FUNCTION OF A > I EDITING

The functions of RNA editing in mammals are still being investigated. On the basis of previously reported evidence a small number of edits alter the coding sequence and activities of certain proteins. An additional small number have direct effects on mRNA splicing, by altering transcript sequence at consensus

splice sites. However, the function of the large majority of RNA edits, which are within intronic or intergenic high copy number repeats, is not known. One possibility is that they have no function at all. They may simply be the collateral damage of an enzyme system which uses dsRNA as a template and which therefore generates large numbers of edits of high copy number repeat elements. According to this hypothesis, the important functional consequences for the cell reside in the small number of coding, splice site and other functional edits. This would be a system of remarkable metabolic profligacy since fewer than 1% (and probably fewer than 0.1%) edits would be functional.

Alternatively, editing of intronic and intergenic high copy number repeats may have a function. One possibility is that RNA editing inhibits non-specific cellular responses to dsRNA which are deleterious to cellular function. These potentially include activation of 2',5'-oligoA synthetase / RNaseL resulting in single stranded RNA degradation, activation of the dsRNA dependent Protein kinase (PKR) resulting in suppression of protein synthesis and activation of the interferon response leading to apoptosis (Kumar and Carmichael, 1998).

Another possibility is that A > I editing prevents gene silencing via the RNAi pathway (Mello and Conte, 2004). It is conceivable that endogenously transcribed dsRNA formed by pairs of inverted Alu repeats are substrates of the dsRNA ribonucleases Dicer, giving rise to Alu derived short interfering RNAs (siRNAs). Given the abundance of Alu sequences in the transcriptome, the number of potential binding sites of Alu siRNAs would be huge and could

have catastrophic effects on the cell. Previous studies in *C. elegans* support the notion that RNA editing abrogates RNAi dependent toxic effects of endogenous dsRNAs (Tonkin and Bass, 2003). An increased number of mismatches generated by editing of dsRNA molecules may limit their deleterious RNAi dependent effects by destabilising the hairpin, by reducing the efficiency of processing (perhaps by retention in the nucleus (Zhang and Carmichael, 2001)), by generating products which are less effective in mediating the effects of RNAi, (for example, by interrupting long, perfectly matched stretches of base pairing) or by other, currently obscure, mechanisms. Our data is broadly consistent with this model, as A > I editing results in an overall increase in the number of mismatches in dsRNA.

An alternative explanation is that dsRNAs formed between inverted Alu repeats are not toxic to the cell, but play a functional role that is regulated by A > I editing. Although closely spaced inverted repeats are apparently toxic to the cell and are underrepresented in the genome (Stenger et al., 2001), our results indicate that nearly 65% of all transcripts have at least one intron with a pair of inverted Alus, and therefore are potential A > I RNA editing substrates. Given that they have accumulated to such a high level in the human genome, it is possible that not all dsRNAs formed by inverted Alu repeats are subject to negative selection.

No function has been ascribed to transcribed inverted repeats in mammals. One possibility, as suggested above, is that they are processed into short RNAs and act in a manner analogous to siRNAs or miRNAs. Rather than

having a toxic effect on cell function, these may be functional molecules which regulate the expression of target transcripts. The role of RNA editing may be to regulate rather than to prevent the entry of Alu derived dsRNA into this pathway.

Interestingly, there were several edited sequences for which, in the simulations, the effect of A > I editing appeared to increase base pairing in dsRNA. This would apparently lead to a small number of dsRNAs becoming more stable and therefore, presumably better substrates for RNAi. Also, A > I RNA editing of a miRNA precursor was recently demonstrated, and predicted to have an effect on the biogenesis and function of the encoded miRNA (Luciano et al., 2004). These results are consistent with a regulatory rather than a preventative role for A > I editing. The use of Alu sequences in such a way may account for their toleration in high abundance in the human genome and in particular their accumulation in gene rich sequences.

There is evidence that A > I RNA editing influences splicing by competing with splicing machinery for RNA at the intron exon junction (Bratt and Ohman, 2003, Flomen et al., 2004), by editing and destroying a branch site adenosine (Beghini et al., 2000) or by creating splice sites (Rueter et al., 1999). In the latter case, a novel splice site is created by ADAR2 editing of an AA dinucleotide in an intronic Alu sequence of its own transcript, to an AG splice site acceptor. In the absence of RNA editing, Alu sequences have been shown to generate splice variants, by virtue of both splicing donor and acceptor consensus sequences within transcribed intronic Alu sequences

(Sorek et al., 2002). The large number of edited intronic Alu sequences identified in this survey includes AA > AI and AG > IG edits. It is therefore possible that regulation of splicing by RNA editing of intronic Alu sequences is widespread. However, none of the edited Alu sequences identified in this survey were spliced, and given that intronic Alu RNA editing substrates are widespread (>60% of all transcripts contain an intronic inverted Alu repeat), it is difficult to envisage specific regulation of splicing through RNA editing.

Whatever the function of A > I editing, it is necessary to account for the observation that the extent of A > I editing and the expression levels of ADAR editing enzymes varies between tissues. It is conceivable that the requirement for RNA editing in a particular tissue is linked to the fate of endogenous dsRNA or the product of dsRNA metabolism in that tissue. For example, cells in which endogenous dsRNA can have deleterious consequences (perhaps by eliciting an RNAi response), may require RNA editing to prevent such a response occurring. Conversely, RNA editing of dsRNA may not be as important in tissues in which endogenous dsRNAs do not have such an effect.

Finally, the association of A > I editing with high copy repeats suggests that A > I editing may function in the biology of retrotransposons. For example, it is possible that A > I editing may lead to the mutation and inactivation of transcribed Alus to prevent their re-insertion into the genome. However, active Alus tend to be transcribed under the control of their own promoters, rather than as components of other transcripts, and therefore would not necessarily be expected to form the types of dsRNA molecules that were found to be

edited in this survey. The potential for single Alu repeats to form dsRNA structures that are substrates for A > I editing is unclear, but seems to be low from our data. If active Alus are subject to modification by A > I editing prior to retrotransposition, evidence for this should be present in the sequence of Alus in the human genome, and may be detectable among other causes of variation such as error prone reverse transcription, and conventional DNA mutation.

There are clearly many interesting unanswered questions regarding the function of RNA editing in human cells. This thesis describes a survey of the patterns of RNA editing in the human brain, and forms a basis for future analyses.