I declare that this thesis is my own account of my research and contains as its main content work that has not previously been submitted for a degree at any tertiary education institution.

Craig Stuart Pace
15 June 2006
ABSTRACT

This thesis addresses two important topics in HIV-1 medicine; (i) the clinical relevance of pre-treatment G→A hypermutation and the contribution of host and viral genetics to its development and; (ii) the influence of genetic variation in host enzymes responsible for antiretroviral drug metabolism on response to therapy. These themes are outlined below.

HIV-1 Hypermutation

At present, limited data exists regarding the relative roles of host encoded cytidine deaminases APOBEC3G and APOBEC3F in promoting G→A hypermutation of HIV-1 proviral DNA in vivo, nor the clinical relevance of hypermutation or the influence of genetic diversity of the APOBEC3G locus and of the viral encoded vif protein that counteracts the action of APOBEC3G. The analyses contained within this thesis demonstrate that within the WA HIV cohort, clinically relevant hypermutation is restricted to a minority of individuals and is mediated predominantly by APOBEC3G. In this study, the presence of HIV-1 hypermutation had a substantially greater effect on plasma viremia than other known host antiviral factors such as CCR5Δ32 or specific HLA-B alleles. Furthermore, the considerable genetic diversity of the vif gene is likely to make a greater contribution to the development of hypermutation than the limited genetic diversity of the APOBEC3G gene in Caucasians. These data indicate that G→A hypermutation is a clinically relevant phenomenon and may provide a fresh perspective to the area of HIV/AIDS therapies.
Genetic Determinant of HIV-1 Treatment Response

Thymidine kinase 2 (TK2) and thymidylate kinase (dTMPK) are rate limiting enzymes for the metabolism of the antiretrovirals d4T and AZT, respectively, and are thus central to the antiviral efficacy and toxicity of these agents. However, the genetic diversity of TK2 and dTMPK and their influence on toxicities associated with their use is largely unknown. The results discussed in this thesis indicate that in contrast to the highly conserved TK2 locus, the dTMPK locus of Caucasian individuals, including regulatory regions potentially influencing transcription and translation, is considerably polymorphic and organised into five common haplotypes.

The results regarding the contribution of dTMPK genetic variation to toxicities associated with AZT therapy are encouraging. A common dTMPK haplotype had significant, albeit modest, effect on haematological parameters (haemoglobin and mean corpuscular volume) in HIV-infected patients, although no AZT-specific treatment effect was observed in this relatively haematologically stable cohort. In addition, another common dTMPK haplotype provided significant protection against AZT-induced adipocyte mtDNA depletion in a pilot study of AZT- and d4T-treated individuals. The dTMPK haplotypes characterised in this thesis should facilitate further studies regarding dTMPK genetic variation in HIV-1 infection and response to treatment, which are warranted from the clinical results presented herein.
TABLE OF CONTENTS

Abstract i
Table of Contents iii
List of Figures vii
List of Tables xi
Abbreviations xii
Acknowledgements xv

CHAPTER ONE - POPULATION LEVEL ANALYSIS OF HYPERMUTATION AND ITS RELATIONSHIP WITH APOBEC3G AND VIF GENETIC VARIATION 1

1.1 General Introduction and Review of the Literature 2

1.1.1 Hypermutation 3
1.1.1.1 Description of HIV-1 Hypermutation 3
1.1.1.2 Cause of HIV-1 Hypermutation 3
1.1.1.3 Mechanism of APOBEC3-Mediated Hypermutation 5

1.1.2 APOLIPOPROTEIN B MRNA EDITING ENZYME CATALYTIC(-LIKE) POLYPEPTIDES 7
1.1.2.1 Introduction to the AID/APOBEC Superfamily 7
1.1.2.2 Structure of AID/APOBEC Proteins 8
1.1.2.3 Evolution of the AID/APOBEC Superfamily 10
1.1.2.4 Expansion of the APOBEC3 Locus 11
1.1.2.5 Expression of APOBEC3 Proteins 16
1.1.2.6 Transcriptional Regulation of APOBEC3G 18
1.1.2.7 Post-Transcriptional Regulation of APOBEC3G 18
1.1.2.8 APOBEC3G-Vif Interaction 21
1.1.2.9 Virion Encapsidation of APOBEC3G 24
1.1.2.10 Antiviral Domains of APOBEC3G 26

1.1.3 Viral Infectivity Factor (Vif) 29
1.1.3.1 HIV Accessory/Regulatory Genes 29
1.1.3.2 Role of Vif 30
1.1.3.3 Vif-Mediated Degradation of APOBEC3G 32
1.1.3.4 Vif Encapsidation 35
1.1.3.5 Regulation of Vif 36
1.1.3.6 Vif-APOBEC3 Binding 37

1.1.4 Clinical Significance 41
1.1.4.1 APOBEC3G Genetic Variation 41
1.1.4.2 Vif Genetic Variation 43

1.1.5 Hypothesis & Aims 45
1.1.5.1 Hypothesis 45
1.1.5.2 Aims 45
1.2 MATERIALS & METHODS 46

1.2.1 PATIENT SELECTION 46
1.2.2 HIV-1 CLADE ASSIGNMENT 46
1.2.3 AMPLIFICATION AND SEQUENCING OF HIV-1 PROVIRAL DNA, MEASUREMENT OF PRE-TREATMENT HIV RNA LEVELS, HLA AND CCR5 GENOTYPING 47
1.2.4 ANALYSIS OF G→A SUBSTITUTIONS 48
1.2.5 ANALYSIS OF APOBEC3G AND APOBEC3F TARGET MOTIFS 49
1.2.6 ANALYSIS OF APOBEC3G- AND APOBEC3F-MEDIATED HYPERMUTATION 50
1.2.7 APOBEC3G ALLELE FREQUENCIES 51
1.2.8 STATISTICAL ANALYSIS 52

1.3 RESULTS 53

1.3.1 ANALYSIS OF G→A SUBSTITUTIONS 53
1.3.2 ANALYSIS OF APOBEC3G AND APOBEC3F DINUCLEOTIDE TARGET MOTIFS 54
1.3.3 CLASSIFICATION OF APOBEC3G- & APOBEC3F-HYPERMUTATED HIV-1 SEQUENCES 56
1.3.4 HIV-1 GENOMIC DISTRIBUTION OF G→A HYPERMUTATION 59
1.3.5 ASSOCIATION OF VIF AMINO ACID POLYMORPHISMS AND G→A HYPERMUTATION 59
1.3.6 APOBEC3G GENETIC VARIATION AND G→A HYPERMUTATION 65
1.3.7 G→A HYPERMUTATION AND HIV-1 VIREMIA 67

1.4 DISCUSSION 68

1.5 SUPPLEMENTARY DATA 73

CHAPTER TWO - CHARACTERISATION OF THE HUMAN DEOXYTHYMIDYLATE KINASE AND THYMIDINE KINASE 2 GENES AND THEIR INFLUENCE ON AZT AND D4T TOXICITIES 79

2.1 GENERAL INTRODUCTION AND REVIEW OF THE LITERATURE 80

2.1.1 AZT AND 4T THERAPY 80
2.1.1.1 Mode of Action of NRTIs 81
2.1.2 CLINICAL TOXICITIES OF THYMIDINE-ANALOGUE THERAPY 83
2.1.2.1 Haematological Toxicities 83
2.1.2.2 Myopathy
2.1.2.3 Lactic Acidosis and Hyperlactataemia
2.1.2.4 Neuropathy
2.1.2.5 Lipoatrophy
2.1.2.6 Pancreatitis
2.1.2.7 Myopathy
2.1.3 MECHANISMS OF THYMIDINE-ANALOGUE INDUCED MITOCHONDRIAL DYSFUNCTION
2.1.3.1 Mitochondrial DNA Depletion
2.1.3.2 Non-Mitochondrial DNA Depletion Mechanisms of AZT Induced Mitochondrial Toxicity
2.1.4 NON-MITOCHONDRIAL DYSFUNCTION MECHANISMS OF AZT TOXICITY
2.1.4.1 Genotoxicity
2.1.4.2 Inhibition of Heme Synthesis
2.1.4.3 Nucleotide Pool Imbalance
2.1.5 PHOSPHORYLATION OF AZT & d4T
2.1.5.1 Thymidine Kinase 2
2.1.5.2 Deoxythymidylate Kinase
2.1.5.3 Nucleoside Diphosphate Kinase
2.1.6 HYPOTHESES & AIMS
2.1.6.1 Hypothesis
2.1.6.2 Aims
2.2 M ATERIALS & M ETHODS
2.2.1 IDENTIFICATION OF SNPS AND CHARACTERISATION OF HUMAN TK2 & dTMPK HAPLOTYPES
2.2.1.1 Amplification of dTMPK
2.2.1.2 Amplification of TK2
2.2.2 THE INFLUENCE OF dTMPK GENETIC VARIATION ON HAEMATOLOGICAL RESPONSE TO AZT THERAPY
2.2.3 THE INFLUENCE OF dTMPK GENETIC VARIATION ON AZT-INDUCED mtDNA DEPLETION
2.2.4 STATISTICAL ANALYSIS
2.3 R ESULTS
2.3.1 IDENTIFICATION OF TK2 AND dTMPK SNPS
2.3.2 CHARACTERISATION OF dTMPK HAPLOTYPES
2.3.3 HAEMATOLOGICAL RESPONSE TO AZT AND d4T
2.3.4 INFLUENCE OF dTMPK GENETIC VARIATION ON HAEMATOLOGICAL RESPONSE TO ZIDOVUDINE
2.3.4.1 Haemoglobin
2.3.4.2 Mean Corpuscular Volume
2.3.5 INFLUENCE OF dTMPK SNPS AND HAPLOTYPES ON AZT-INDUCED mtDNA DEPLETION
2.4 Discussion 133

3 References 138
**LIST OF FIGURES**

| Figure 1.1 | Replication of HIV-1 in non-permissive and permissive T cells. | 5 |
| Figure 1.2 | Schematic representation of the mechanism of APOBEC3G mediated G→A hypermutation. | 6 |
| Figure 1.3 | Domain organisation of AID/APOBEC proteins. | 9 |
| Figure 1.4 | Structure of the cytidine deaminase catalytic domain. | 10 |
| Figure 1.5 | APOBEC3 locus on chromosome 22. | 12 |
| Figure 1.6 | Hypothetical scheme of the origin and divergence of the APOBEC3 family in mammals. | 13 |
| Figure 1.7 | Phylogeny of APOBEC3G in Primates. | 15 |
| Figure 1.8 | APOBEC3G colocalises with mitochondria. | 16 |
| Figure 1.9 | Northern blot analysis of APOBEC3 genes. | 17 |
| Figure 1.10 | APOBEC3G mRNA levels are regulated by PKC and MEK1/2. | 19 |
| Figure 1.11 | Inducibility of HMM APOBEC3G complex formation in primary CD4+ T cells with different stimuli, and correlation of the HMM complex with permissivity for HIV infection. | 20 |
| Figure 1.12 | The APOBEC3G D128K mutation selectively allows APOBEC3G to bind HIV-1 *vif* or SIVagm *vif*. | 22 |
| Figure 1.13 | HIV-1 *Vif* is able to interact and form complexes with APOBEC3G D128K mutant (A) but fails to reduce steady-state levels of the mutant APOBEC3G protein (B). | 23 |
| Figure 1.14 | Mutations of the zinc coordinating (Cys97 and Cys100, 1CCAA) and RNA binding (Phe70 and Tyr91, 1FYAA) residues of CD1 but not of the CD2 (2CCAA & 2FFAA) domain reduces HIV-1 encapsidation of APOBEC3G. | 25 |
| Figure 1.15 | The Cytidine Deaminase Activity of APOBEC3G Is Not Required for an Antiviral Effect. | 27 |
Figure 1.16 Schematic representation of HIV-1 genome.

Figure 1.17 Replication of vif-positive and vif-negative HXB2 viruses in CD4+ human T-cell lines.

Figure 1.18 HIV-1 vif interacts with cellular proteins Cul5, ElonginB and ElonginC (A) to promote the polyubiquitination and subsequent proteasomal-mediated degradation of vif-APOBEC3G complexes (B).

Figure 1.19 Vif C114S and C133S mutants are able to bind Cul5 (A) but fail to ubiquitinate APOBEC3G (B) despite similar levels of vif, EloBC and Cul5 expression (C).

Figure 1.20 Phosphorylation of HIV-1 proteins by MAPK.

Figure 1.21 Stable coexpression of APOBEC3G and vif.

Figure 1.22 Vif peptides comprised of vif amino acids 85-99 and 169 to 192 significantly inhibit interaction between phage-displayed APOBEC3G and vif.

Figure 1.23 The APOBEC3G 186R variant associated with accelerated progression to AIDS exhibits similar antiviral activity as the 186H variant.

Figure 1.24 G→A burden and G→A preference are highly correlated.

Figure 1.25 G→A hypermutation occurred predominantly in the dinucleotide sequence contexts targeted by APOBEC3G (GG) and, to a lesser extent, APOBEC3F (GA).

Figure 1.26 The consolidated 3G scores reflect a bimodal distribution.

Figure 1.27 Non-specific G→A hypermutation (A), APOBEC3G-specific G→A substitutions (B) and APOBEC3G-mediated hypermutation (C) scores are relatively constant along the HIV-1 genome.

Figure 1.28 Consensus vif amino acid sequence and polymorphism frequency among non-hypermutated sequences.

Figure 1.29 Schematic representation of HM and truncated vif proteins identified in vivo from putative translation of proviral DNA sequences.
Figure 1.30S  Binding sites of oligonucleotide primers used for the amplification (black) and sequencing (blue and red) of the entire human APOBEC3G gene, including 2kb upstream of the transcription binding site and 0.5kb 3’ of the 3’UTR.

Figure 1.31S  Chromatograph results of sequencing pooled PCR. products homozygous for alternative alleles at varying proportions to estimate the lower limit of allele detection by sequencing.

Figure 1.32S  Power calculations used to determine sample size required to detect significant differences in APOBEC3G allele frequencies between HM-3G and NH sequences based on the allele frequencies of the HM-3G sequences and those estimated from a pooled DNA control sample.

Figure 2.1  Structure of nucleoside reverse transcriptase inhibitors and their corresponding deoxynucleoside.

Figure 2.2  Focal cytochrome oxidase deficiency is a unique and consistent feature of AZT-associated myopathy.

Figure 2.3  Rise in population average venous lactate concentrations after start of highly active antiretroviral therapy (HAART) in treated patients stavudine (dashed line) and zidovudine (solid line).

Figure 2.4  Proportion of patients with peripheral fat depletion according to each body area, in the zidovudine and stavudine arms.

Figure 2.5  The effect of AZT on mtDNA/ nDNA ratio in vitro and in vivo.

Figure 2.6  The effect of AZT on cell growth and haemoglobin production.

Figure 2.7  TMPK is at the junction of the de novo and salvage pathways of dTTP synthesis.

Figure 2.8  Comparison of phosphorylation of 2 µM D4T with 2 µM AZT in CEM cells.

Figure 2.9  The relationship between MTT cytotoxicity (bars) and AZT-MP and AZT-TP (trend lines) in five cell lines.
Figure 2.10  The reduced phosphorylation of nucleoside analogues is due to the absence of the 3’-OH group which reduces the catalytic efficiency rather than the binding affinity.

Figure 2.11  Longitudinal mean corpuscular volume (MCV) profiles of AZT-recipients

Figure 2.12  Longitudinal haemoglobin concentration profiles of individuals initiating AZT therapy

Figure 2.13  Schematic representation of the human dTMPK gene and location of identified SNPs.

Figure 2.14  Schematic representation of the human TK2 gene and location of identified SNPs.

Figure 2.15  dTMPK haplotype pedigrees.

Figure 2.16  AZT recipients homozygous for 3’ haplotype B have significantly higher on-treatment haemoglobin concentrations (A) and significantly less likely to have on-treatment haemoglobin concentrations less than 130 g/L (B).

Figure 2.17  No association existed between the 3’ haplotype B and on-treatment haemoglobin concentrations in d4T-recipients.

Figure 2.18  Mean Δ haemoglobin concentrations (A) and the proportion of individuals that experienced an increase in haemoglobin concentrations in response to AZT-therapy (B) were similar in individuals regardless of homozygosity for the 3’haplotype B or haplotype 3, respectively.

Figure 2.19  AZT recipients homozygous for the dTMPK haplotype 1 have significantly higher adipocyte mtDNA values than non-homozygous AZT recipients (A), despite similar duration of AZT exposure (B).

Figure 2.20  mtDNA values in AZT recipients homozygous for the dTMPK haplotype 1 reflect that observed in recipients of non-thymidine-analogue therapy, HIV-1+ individuals not receiving antiretroviral therapy or HIV-1 seronegative controls.
# List of Tables

| Table 1.1 | Patient and sequence characteristics of patients harbouring putative APOBEC3G-, APOBEC3F-hypermutated and non-hypermutated HIV-1 proviral DNA. | 58 |
| Table 1.2 | APOBEC3G allele frequencies among 8 patients with marked G→A hypermutation. Comparisons with a pooled DNA control sample and a published study. | 66 |
| Table 1.3S | Oligonucleotide sequences used for the amplification and sequencing of the entire human APOBEC3G gene, including 2.0 kb 5’ of the transcription start site and 1 kb 3’ of the transcription stop site. | 73 |
| Table 2.1 | Oligonucleotide characteristics used for the amplification and sequencing of the human dTMPK gene. | 111 |
| Table 2.2 | Oligonucleotide characteristics used for the amplification and sequencing of the human TK2 gene. | 112 |
| Table 2.3 | Human deoxythymidylate kinase haplotypes identified from studies of family members. | 121 |
| Table 2.4 | Characteristics of AZT and d4T recipients. | 122 |
ABBREVIATIONS

3TC    lamivudine
A      adenine
AACTG  Adult AIDS Clinical Trials Group
AAN    antiretroviral-associated neuropathy
ABC    abacavir
ADP    adenosine diphosphate
agm    African green monkey
AID    activation-induced deaminase
AIDS   acquired immunodeficiency syndrome
Ala    alanine
AMT    3’-amino-3’deoxythymidine
ApoB   apolipoprotein B
APOBEC apolipoprotein B mRNA-editing enzyme
Arg    arginine
Asn    asparagine
ATP    adenosine triphosphate
AZT    zidovudine
C      cytosine
CCR    chemokine receptor
CD     cytidine deaminase
CD4    cluster of differentiation maker 4
cDNA   complimentary deoxyribose nucleic acid
CFU-GM colony forming unit-granulocyte macrophage
C-terminal carboxy-terminal
Cul    cullin
Cys    cysteine
D      aspartic acid
d4T    stavudine
da     deoxyadenine
dc     deoxycytidine
dCTP   deoxycytidine triphosphate
ddC    zalcitabine
ddI    didanosine
dG     deoxyguanidine
DNA    deoxyribose nucleic acid
dNTP   deoxynucleoside triphosphate
DP     diphosphate
dT     deoxythymidine
dTMPK  deoxthymidylate kinase
dTTP   deoxythymidine triphosphate
E      glutamic acid
Elo    elongin
ERK    extracellular protein kinase
F      phenylalanine
FDA    Federal Drug Administration
fl     femtolitres
FLT    3’-fluoro-3’-deoxythymidine
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe</td>
<td>phenylalanine</td>
</tr>
<tr>
<td>Pi</td>
<td>inorganic phosphate</td>
</tr>
<tr>
<td>Pro</td>
<td>proline</td>
</tr>
<tr>
<td>Q</td>
<td>glutamine</td>
</tr>
<tr>
<td>R</td>
<td>arginine</td>
</tr>
<tr>
<td>Rev</td>
<td>regulator of viral protein synthesis</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RT</td>
<td>reverse transcriptase</td>
</tr>
<tr>
<td>S</td>
<td>serine</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>Ser</td>
<td>serine</td>
</tr>
<tr>
<td>SIV</td>
<td>simian immunodeficiency virus</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
<tr>
<td>SOCS</td>
<td>suppressor of cytokine signalling</td>
</tr>
<tr>
<td>T</td>
<td>thymidine</td>
</tr>
<tr>
<td>Tat</td>
<td>transactivating transcription factor</td>
</tr>
<tr>
<td>TDF</td>
<td>tenofovir</td>
</tr>
<tr>
<td>Thr</td>
<td>threonine</td>
</tr>
<tr>
<td>TK</td>
<td>thymidine kinase</td>
</tr>
<tr>
<td>TNF</td>
<td>tumour necrosis factor</td>
</tr>
<tr>
<td>TP</td>
<td>triphosphate</td>
</tr>
<tr>
<td>TSS</td>
<td>transcription start site</td>
</tr>
<tr>
<td>Tyr</td>
<td>tyrosine</td>
</tr>
<tr>
<td>U</td>
<td>uridine</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>Vif</td>
<td>viral infectivity factor</td>
</tr>
<tr>
<td>Vpr</td>
<td>viral protein R</td>
</tr>
<tr>
<td>Vpu</td>
<td>viral protein U</td>
</tr>
<tr>
<td>wt</td>
<td>wild type</td>
</tr>
<tr>
<td>Y</td>
<td>tyrosine</td>
</tr>
</tbody>
</table>
Acknowledgements

Firstly, I would like to thank Professors Simon Mallal and Ian James for giving me the opportunity to study for a PhD degree. Without their support, funding and most of all faith in me to complete, firstly an Honours degree, and then a PhD degree, I would not be in this position today. In that regard, I would also like to thank Neill Hodgen, who was instrumental in me obtaining a position with the Centre when I first finished my undergraduate studies. I would also like to thank my supervisor, Dr David Nolan, for his encouragement, advice and invaluable time; no matter how busy he was, he was always able to make the time to discuss any aspect of my thesis.

My gratitude also extends to Jean Keller for performing the database queries required for the hypermutation analyses and statistical advice, Dr Silvana Gaudieri for general methodology issues, Dr Larry Park for HIV-1 clade assignment, Dr Corey Moore for establishing the HIV-1 sequence database and Filipa Carvalho for performing a large proportion of the HIV-1 sequencing.

I would also like to thank my fellow S blockers; PhD students Nicole Prada, Coral-Ann Almeida and Bree Foley; computer staff Mark Shaw and Iain Bradley and Niamh Keene for thesis advice. Finally, I’d like to thank my family for their encouragement.