1-1-2014

Genetic And Biochemical Studies Of Human Apobec3h Enzyme

Thisari Sachithra Aluthgama Guruge
Wayne State University,

GENETIC AND BIOCHEMICAL STUDIES OF HUMAN APOBEC3H ENZYME

by

THISARI SACHITHRA ALUTHGAMA GURUGE

THESIS

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

2014

MAJOR: CHEMISTRY (Biochemistry)

Approved by:

_____

Advisor Date

DEDICATION

To my family

ACKNOWLEDGMENTS

First and foremost, I would like to convey my heartfelt gratitude to my thesis advisor, Dr. Ashok S. Bhagwat for the continuous support and advices he has given to me during my research career. I am truly grateful to you for the encouragement and immense support throughout the past three years.

I would like to thank my thesis committee, Dr. Tamara L. Hendrickson and Dr. Mary Kay Pflum for their comments, suggestions and presence.

I thank the Department of Chemistry, Wayne State University, for the support given to make this research project a success.

I would like to convey my gratitude to Dr. Priyanga Wijesinghe for training me when I first joined the lab and also for his encouraging words. I would also like to thank Dr. Ashoka Shyamalee Kandegedara for constructive comments and suggestions on protein purification steps. A special thank goes to Amit Kumar for assisting me in using the FPLC system.

I convey my heartfelt gratitude to all the friends who helped me during my time at Wayne State University. I thank Dilani Gigiriwala Gamage, Dr. Gayathri Silva, Thiloka Ariyasena and Dhanusha Nalawansha for being there for me during good times as well as bad times. Your support was invaluable.

I would also like to thank my past and present lab members for the times we shared as group. I thank Sophia Shalhout for training me on mammalian cell culture work. I would also like to thank Shanqiao Wei, Vimukthi Asgiriya Senevirathne, Sachini Siriwardena, Dr. Anjali Patwardhan, Dr. Marina Bakhmutsky and Anita Chalasani for the stimulating discussions and support given to me during my research career.

I would also like to thank the administrative staff in the Department of Chemistry, Wayne State University. Melissa Barton, Deborah McCreless, Bernadette Miesik and Diane Kudla assisted me with any question, whenever I needed.

My sincere gratitude goes to the Sri Lankan Students Association (SLSA) for the support given to me from the time of arrival to the U.S. and helping me to adjust the life here. I would like to convey my sincere gratitude to all the other people who helped me throughout my career in Wayne State University.

I am really thankful to my beloved parents for their invaluable support and encouragement and for paving the right path for me. Words cannot explain how grateful I am to both of you!

Last but not least, I would like to thank my husband, for being there for me in all times. Thank you so much for your motivation and support given to me when I needed it the most!

TABLE OF CONTENTS

DEDICATION ---- ii

ACKNOWLEDGMENTS ---- iii

LIST OF TABLES ---- vii

LIST OF FIGURES ---- viii

LIST OF ABBREVIATIONS ---- ix

CHAPTER 1: INTRODUCTION ---- 1

1.1 The AID/APOBEC family ---- 1

1.2 APOBEC3 sub-family ---- 2

1.3 Mouse APOBEC3 enzyme ---- 3

1.4 Human APOBEC3 enzymes ---- 4

1.5 HIV-1 restriction by APOBEC3 proteins ---- 6

1.6 APOBEC3H enzymes ---- 8

CHAPTER 2: MATERIALS AND METHODS ---- 21

2.1 Cloning of human APOBEC3H ---- 21

2.2 Purification of GST-APOBEC2H-200 ---- 24

2.3 Western Blotting of GSTA3H200 ---- 26

2.4 Kanamycin resistance reversion assay ---- 27

2.5 Purification of poly-histidine tagged A3H-183 using nickel affinity column ---- 28

2.6 Western Blotting of his-A3H183 ---- 30

2.7 Size exclusion chromatography of poly-histidine tagged A3H-183 ---- 31

2.8 *In vitro* activity assays of purified APOBEC3H proteins ---- 31

CHAPTER 3: RESULTS AND DISCUSSION ---- 33

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS ---------- 43

REFERENCES ---------- 44

ABSTRACT ---------- 52

AUTOBIOGRAPHICAL STATEMENT ---------- 53

LIST OF TABLES

Table 1: Summary of activity and distribution of human APOBEC3H ---------- 19

Table 2: The oligonucleotide sequences. ---------- 22

LIST OF FIGURES

Figure 1: Members of the AID/APOBEC family of enzymes in higher vertebrates. 1

Figure 2: The arrangement of the APOBEC3 genes in human chromosome 22. 3

Figure 3: Mechanism of retroviral restriction mediated by APOBEC3G enzyme. 7

Figure 4: Representation of the four major splice variants of human APOBEC3H. 13

Figure 5: Human APOBEC3H splice variants from ENSEMBL genome browser 14

Figure 6: Four splice variants of human APOBEC3H found in PBMCs 14

Figure 7: pSU24 plasmid map and MCS. 23

Figure 8: pGEX4T3 plasmid map and MCS. 23

Figure 9: pET28a+ plasmid map and MCS. 24

Figure 10: The basis of the kanamycin resistance reversion assay. 28

Figure 11: Kanamycin resistance reversion assay results for pSUA3H-200. 33

Figure 12: 12% SDS-PAGE gel for GSTA3H200 purification. 34

Figure 13: Western blot for GSTA3H200 35

Figure 14: DNA cytosine deaminase activity assay for partially purified GSTA3H200 35

Figure 15: Kanamycin resistance reversion assay results for pSUA3H183. 36

Figure 16: 15% SDS-PAGE gel for nickel affinity purified APOBEC3H-183. 37

Figure 17: Western blot for protein fractions obtained in nickel affinity purification. 38

Figure 18: Biochemical activity assay for partially purified APOBEC3H-183. 39

Figure 19: Biochemical activity assay for partially purified APOBEC3H-183. 40

Figure 20: Size exclusion chromatogram. 41

Figure 21: SDS-PAGE gel (15%) with APOBEC3H-183 from a size exclusion column. 41

Figure 22: DNA cytosine deaminase activity assay of his-APOBEC3H-183. 42

LIST OF ABBREVIATIONS

6-FAM: 6-carboxyfluorecein

A: Adenine

A3A: APOBEC3A

A3H: APOBEC3H

agm: African Green Monkey

AID: Activation Induced Deaminase

APOBEC: **Apo**lipoprotein **B** mRNA editing **E**nzyme- **C**atalytic polypeptide

C: Cytosine

CDD: Cytosine Deaminase Domain

cDNA: Complementary Deoxyribonucleic Acid

CTD: Caboxy Terminal Domain

Cys: Cysteine

DTT: Dithiothreitol

E.coli : *Escherichia coli*

EDTA: Ethylenediaminetetraacetic acid

G: Guanine

Glu: Glutamic acid

GST: Glutathione S-transferase

Hap: Haplotype

HBV: Hepatitis B Virus

HCl: Hydrochloric acid

His: Histidine

HIV-1: Human Immunodeficiency Virus-1

HPV: Human Papilloma Virus

IPTG: Isopropyl β-D-1-thiogalactopyranoside

Kan^R: Kanamycin Resistant

LB: Luria Broth

LINE: Long Interspersed Nuclear Elements

MCS: Multiple Cloning Site

MLV: Murine Leukemia Virus

mRNA: Messenger Ribonucleic acid

mug: Mismatch Uracil Glycosylase

NaCl: Sodium Chloride

NaOH: Sodium Hydroxide

NES: Nuclear Export Signal

NMR: Nuclear Magnetic Resonance

NTD : N Terminal Domain

OWM: Old World Monkey

PAGE: Polyacryamide Gel Electrophoresis

PBMC: Peripheral Blood Mononuclear Cells

Pro: Proline

PVDF: Polyvinylidenefluoride

RTase: Reverse Transcriptase

SDS: Sodium Dodecyl Sulfate

SIV: Simian Immunodeficiency Virus

SNP: Single Nucleotide Polymorphism

sv: Splice Variant

T: Thymine

TBE: Tris Borate EDTA

U: Uracil

UDG: Uracil DNA Glycosylase

Vif: Virulence Infectivity Factor

CHAPTER 1: INTRODUCTION

1.1 The AID/APOBEC family

The AID and APOBEC family of proteins catalyze the deamination of cytosines to uracils in single stranded DNA and/or RNA. They all share a conserved zinc binding domain, His-X-Glu-X_{25-31}-Pro-Cys-X_{2-4}-Cys. (X is any amino acid). The zinc (Z)-coordinating motif is important in cytosine deamination. They possess either one or two of these domains (highlighted in Figure 1). Humans have seven APOBEC3 members.

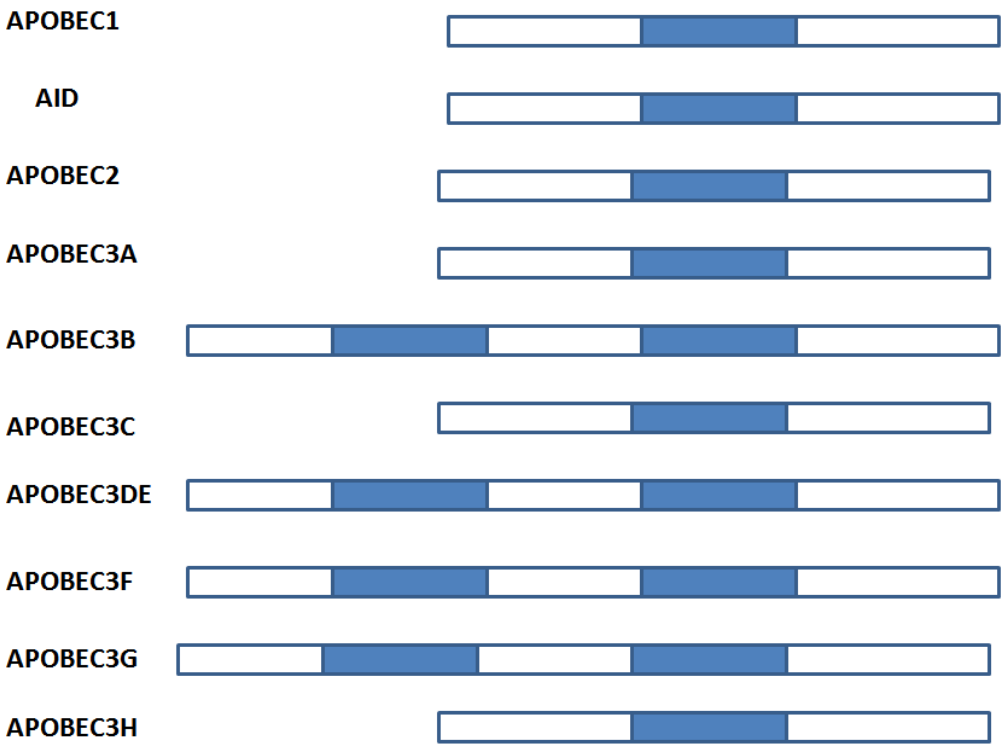

Figure 1: Members of the AID/APOBEC family of enzymes in higher vertebrates. The zinc-binding domain is highlighted.

APOBEC1 was the first enzyme to be identified in this family of enzymes, which plays an essential role in lipid metabolism [1, 2]. APOBEC1 introduces a premature translation termination codon in the pre-mRNA of Apolipoprotein B100, producing a shorter version, namely, Apolipoprotein B48. The two versions have differential activities regulating lipoproteins in liver. Mammalian APOBEC1 acts on RNA, but as discovered by Harris *et al*. in 2002, it can also act on single stranded DNA [3].

AID was the second member to be identified in this family by Muramatsu *et al.* in 1999 [4]. It deaminates cytosines in single stranded DNA [3, 5]. AID plays a remarkable role in adaptive immunity in higher vertebrates. Specifically, AID is essential for antibody diversification and for the genearation of high affinity antibodies. In B cell immunoglobulin genes, AID introduced uracils are processed in different DNA repair and recombination pathways leading to class switch recombination and somatic hypermutation [6, 7]. These processes are essential in efficient immune defense against pathogens.

1.2 APOBEC3 sub-family

Over the past 100 million years, APOBEC enzymes have evolved as a protective mechanism of immunity to prevent viruses and pathogens from infecting the organisms. The ancestral zinc binding deaminase domain has duplicated several times in humans within 150 kb in chromosome 22. This duplication encodes several APOBEC3 proteins [8]. The human APOBEC3 family consists of seven members, containing one or two zinc binding deaminase domains, namely, APOBEC3A-APOBEC3H. Primates also have multiple APOBEC3 genes that have undergone expansion and duplication, while rodents have a single APOBEC3 gene in their genomes [9]. The importance of this expansion of the APOBEC3 genes may be implicated in the protective function of restricting evolving

endogenous and exogenous retroviruses or retroelements in these higher organisms. More specifically, APOBEC3 enzymes are potent inhibitors of replication of exogenous retroviruses such as HIV-1 [10] although with varying levels [11, 12].

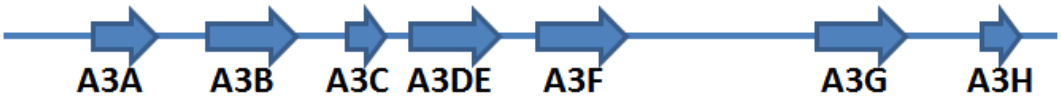

Figure 2: The arrangement of the APOBEC3 genes in human chromosome 22.

1.3 Mouse APOBEC3 enzyme

In contrast to humans, mice have one APOBEC3 gene. It is located in chromosome 15 which is syntenic to human chromosome 22 [13]. This gene encodes for a double domain DNA cytosine deaminase. In contrast to human APOBEC3G, where only the CTD is catalytically active, the NTD is catalytically active in mouse APOBEC3. The NTD is thought to be necessary for encapsidation of this enzyme into retroviruses. It shares ~33% sequence similarity with human APOBEC3G [14, 15]. Mouse APOBEC3 is a powerful restrictor of vif-deficient HIV-1, and the efficiency of restriction is comparable to that of the human APOBEC3G. Mouse APOBEC3 can introduce high levels of G to A mutations in the HIV-1 genome; hence the mechanism of restriction is deaminase dependent [16]. On the other hand, its ability to restrict murine retroviruses, such as MLV is limited compared to human APOBEC3G [17]. In biochemical terms, purified mouse APOBEC3 is capable of deaminating cytosine in short single stranded DNA, preferably towards the 3' end [16]. Mouse APOBEC3 is expressed widely in lymphoid tissue. Despite the useful role played by mouse APOBEC3, it is not essential for the survival or development of mice [18].

1.4 Human APOBEC3 enzymes

APOBEC3A is a single domain DNA cytosine deaminase, which is implicated in its potential for restricting a myriad of retroviruses including HIV-1 [19, 20]. APOBEC3A has also has the potential to restrict certain DNA viruses such as HPV [21] and parvovirus [22]. Overexpressed APOBEC3A is capable of translocating to the nucleus and inducing genomic DNA damage [23]. However, Harris *et al.* suggested that endogenous APOBEC3A is cytoplasmic and incapable of genomic DNA editing [24]. Recently, it was also suggested that APOBEC3A can act as a foreign DNA restricting enzyme in mammalian cells [25]. In addition to DNA cytosine deaminase activity, APOBEC3A is efficient in deaminating 5-methylcytosines, suggesting a possible role in epigenetic gene regulation [26, 27]. APOBEC3A is a myeloid specific enzyme and type-1 interferon can increase its expression 10-1000 fold [19], consistent with its antiretroviral activity. APOBEC3B contains two conserved zinc binding cytosine deaminase domains. Several recent studies have shown that it is an enzymatic source of mutations in several human cancers, including breast cancer [28, 29]. It is the only APOBEC3 enzyme with nuclear localization [29]. Although there is no crystal or NMR structures available, APOBEC3B has been studied extensively as it can be targeted by anticancer drugs. APOBEC3C contains one zinc binding cytosine deaminase domain, and is implicated in HBV [30, 31] and SIV [32] restriction. A high resolution crystal structure of full length APOBEC3C was determined recently [33]. APOBEC3D, which was previously termed, APOBEC3DE is a moderately potent HIV-1 inhibitor in T cells. In contrast, it is a potent inhibitor of Alu retrotransposition. As discussed below, APOBEC3D evolved under positive selection with loss of function mutations [34]. APOBEC3F contains two zinc binding domains. A high resolution crystal structure of the catalytically active APOBEC3F CTD is available [35].

APOBEC3F is capable of restricting vif deficient HIV-1 but it is susceptible to vif mediated degradation [36, 37]. Vif is a HIV-1 encoded viral accessory protein that binds APOBEC3 proteins and targets them for degradation, through ubiquitin proteasome pathway. APOBEC3G is the most studied and well characterized member of the APOBEC3 family of enzymes. The crystal structure of the catalytically active CTD of APOBEC3G was first established by Holden *et a.l* in 2008 [38]. The protein was purified from *E.coli* as a GST fusion. The NMR structure of the CTD of APOBEC3G was also published, but in a more soluble mutant form (2K3A) [39]. Numerous studies show that APOBEC3G possess very efficient HIV-1 restriction potential, but similar to APOBEC3F, it is susceptible to vif-mediated degradation [40, 41].

The structures of the APOBEC3 deaminase domains provide insight into the mechanisms and the activity of the APOBEC3 enzymes and the substrate specificity of these proteins. APOBEC3 enzymes are emerging as targets of potential anti-cancer and anti-retroviral drugs. Furthermore, the residues of APOBEC3 enzymes that are involved in HIV-1 vif binding need to be studied. In addition, to study the mechanism underlying retroviral restriction by APOBEC3 proteins and to understand their role in innate immunity, high resolution crystal or NMR structures would be useful. Unfortunately, some of these enzymes are insoluble and tend to precipitate in high concentrations as the molecules tend to aggregate [42]. Up to now, several groups were able to obtain the structures of the APOBEC3G CTD [39, 42], APOBEC3C [33], APOBEC3A [43] and C-terminal domain of APOBEC3F [35]. APOBEC3H was the focus of my studies is described below.

1.5 HIV-1 restriction by APOBEC3 proteins

To explain the mechanism of HIV-1 restriction by APOBEC3 enzymes, APOBEC3G is taken as the example. Vif-deficient HIV-1 virus infects human CD4+ T cells. APOBEC3G is expressed exclusively in the cytoplasm of these cells [41, 44]. To restrict HIV-1 replication and infection, APOBEC3G is first packaged into the viral core [41]. Once the virus particle infects a new target cell, APOBEC3G plays its role in viral restriction. HIV-1 begins to reverse transcribe the RNA genome to the cDNA intermediate which is the target for APOBEC3G [45]. The enzyme deaminates cytosines to uracils in this minus strand cDNA, which leads to degradation of the product. On the other hand, the uracil-rich cDNA undergoes synthesis of the second DNA strand (plus strand), generating G to A hypermutations in the pro-viral genome, which is detrimental for the replication of the virus [41, 45]. Virally encoded vif protein is capable of interacting with APOBEC3G and targeting it to ubiquitin proteasome mediated degradation [40].

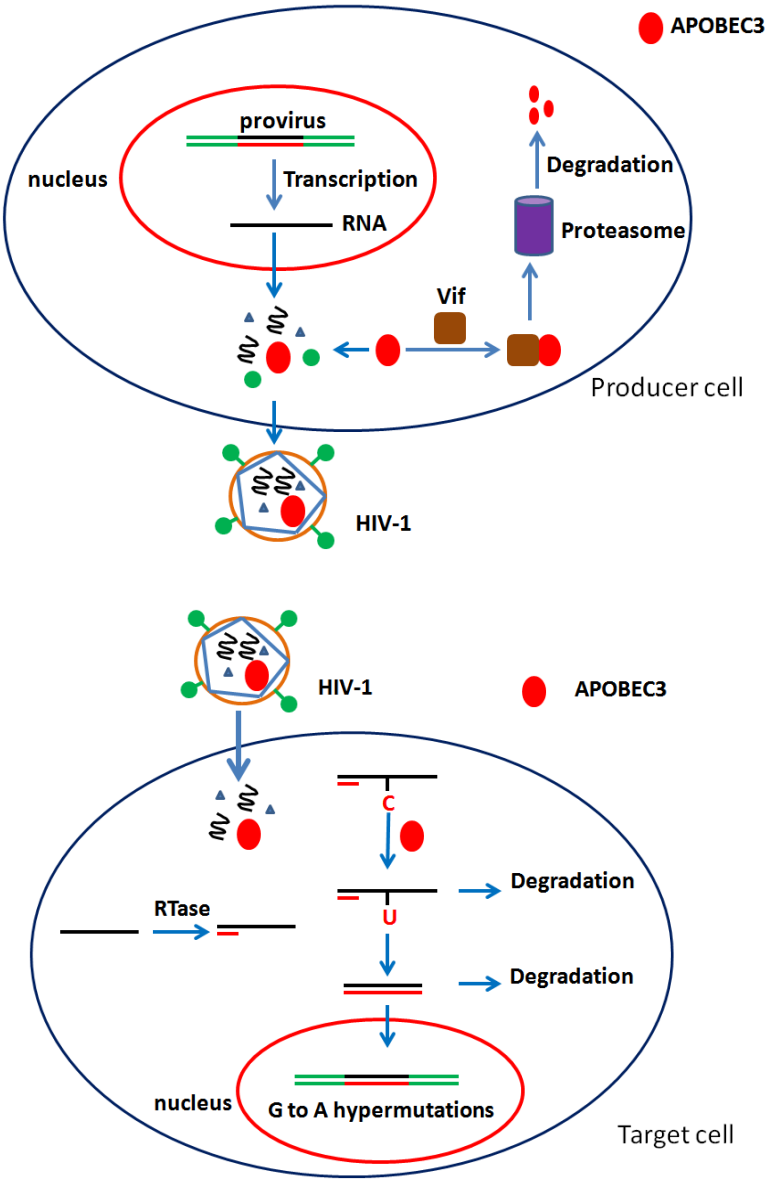

Figure 3: Mechanism of retroviral restriction mediated by APOBEC3 enzyme.

1.6 APOBEC3H enzymes

During host and pathogen protein interactions, evolution plays a key role in domination of either the host or the pathogen. This phenomenon is observed in the evolution of APOBEC3 proteins [46]. As APOBEC3 proteins are involved in restricting retroviruses, the evolution of the viral proteins that counteract the action of APOBEC3 also show strong positive selection [47]. At the same time, host defensive proteins, such as APOBEC3 enzymes, have undergone multiple diversifications and duplications to better counteract viral replication in cells. APOBEC3H evolved from ancestral genes at least 33 million years ago during primate evolution [48].

The human APOBEC3H gene is located downstream of APOBEC3G. It contains a single zinc binding domain. The function of APOBEC3H is poorly characterized. The presence of conserved a cytosine deaminase domain suggested that APOBEC3H might play a role in genome defense. It shares much less amino acid sequence similarity to other single domain cytosine deaminase, APOBEC3A (~29.6%) and APOBEC3C (~34.6%). APOBEC3H mRNA was detected in peripheral blood mononucleated cells (PBMC), testis, ovary and skin [48]. Also, the mRNA level of human APOBEC3H was upregulated >22-fold upon stimulation with interferon in CD4+ T cells [49]. Primate APOBEC3H consists of an evolutionarily distinct cytosine deaminase domain [9]. It is considered as the most evolutionarily divergent APOBEC3 protein and evolved under strong positive selection. Previously it was not identified as a retroviral restriction factor in primates, but the position of the APOBEC3H gene proximal to APOBEC3G and the possession of the conserved zinc binding deaminase domain, lead the researchers to investigate its potential in innate immunity and for the retroviral restriction. Until the whole genome sequencing project was completed, APOBEC3H was not identified as a cytidine

deaminase [9, 13]. APOBEC3H of OWMs, such as macaques, possess strong restriction activity against primate lentiviruses, for example, HIV-1 and SIV [48]. In contrast, the protein expression of human APOBEC3H was poor in primate cells and shows much less restriction ability of retroviruses compared to OWM APOBEC3H [48]. Of note, human APOBEC3H was unable to restrict vif-deficient HIV-1 [48]. In contrast, feline APOBEC3H is a potent inhibitor of HIV-1 and induced G to A mutations in HIV-1 genome [50]. Despite the low levels of human APOBEC3H proteins detected in the transfected cells, mRNA levels were comparable to the level detected in OWM. By successfully expressing OWM and human APOBEC3H in bacteria, OhAinle M. *et al.* [51] demonstrated that both proteins show DNA cytosine deaminase activity. Using a bacterial rifampicin resistance assay, human APOBEC3H was shown to cause ~6.1-fold increase in mutation frequency, while macaque APOBEC3H caused a ~3.1-fold increase over background. By sequencing the *rpoB* gene, a significant increase of G to A mutations was detected, confirming both human and macaque APOBEC3H are proficient in deaminating cytosines in DNA. In contrast, mammalian cells overexpressing human APOBEC3H were used to obtain HIV-1 proviruses, the viral genome did not carry any detectable G to A mutations; however a significant hypermutation was observed in viral genomes obtained from cells expressing macaque APOBEC3H and human APOBEC3G [48].

The weak HIV-1 restriction potential of human APOBEC3H was attributed to its poor expression in primate cells, although it is an active deaminase. Evidence for its potential to mutate the genome of DNA viruses, such as HBV [31] and HPV [21] suggest that human APOBEC3H is active in mammalian cells. In order to investigate the underlying cause for the reduced HIV-1 restriction potential and poor expression of human APOBEC3H in cell culture, Dang Y. *et a.l* [52] compared the amino acid sequences of

several primate APOBEC3H enzymes. Macaque APOBEC3H is 210 amino acids long. The presence of a premature termination codon in the last (fifth) exon in both human and chimpanzee APOBEC3H results in a truncated version in these organisms [52]. The premature termination codon in human APOBEC3H causes a 29 amino acid deletion from the C-terminal end. Loss of these amino acids was attributed to the significantly lower levels of this protein in cells [52]. This deletion was not seen in APOBEC3H genes of macaque and African green monkey. The mRNA level of this shortened APOBEC3H was reduced drastically even though the protein stability was unaffected [52]. Interestingly, restoration of the truncation by repairing the termination codon or increasing the expression of APOBEC3H under the control of a strong promoter, showed increased protein expression. The restored or optimally expressed APOBEC3H showed potent anti HIV-1 and SIV restriction potential [52]. In addition, the authors propose that HIV-1 vif was capable of counteracting the restriction, only when APOBEC3H was expressed well. The barely detectable levels of G to A hypermutations in the HIV-1 genome suggests that although human APOBEC3H is capable of restricting HIV-1, the mechanism is deamination independent. [52].

To further investigate the reasons for the loss of anti-retroviral activity of human APOBEC3H, OhAinle M. *et al.* [51] analyzed the evolution of the activity of hominoid APOBEC3H enzymes. The reconstructed ancestral human APOBEC3H sequence, which showed potent antiretroviral activity, was lost due to two independent polymorphisms that destabilized the protein [51]. The amino acid mutations R105G or ΔN15 contributed to decrease the half life of the protein. Although the majority of the human population showed a reduced amount of stable APOBEC3H expression, hence no antiretroviral activity, individuals of African descent stably express APOBEC3H. Cells from African

populations are capable of inhibiting HIV-1 and LINE-1 retrotransposition [51]. Similar to the ancestral APOBEC3H, this allele does not contain the destabilizing mutations and is sensitive to HIV-1 vif. This suggests that the evolution of human APOBEC3H is under strong negative correlation with its potential to restrict retroviruses, presumably because of the evolution of many other retroviral restriction factors in human cells or more functionally proficient APOBEC3 enzymes. As explained below, the population specific variation in the stability of human APOBEC3H is identified due to the presence of multiple substitution mutations in the gene in different human lineages.

Interestingly, multiple polymorphisms exist naturally in human APOBEC3H. Primarily, the G105R, K121E and E178D cluster of substitution mutations are observed in African populations frequently, but are considerably less common in Asian and European populations [53]. This cluster, not the wild type, has shown restriction potential against HIV-1, Alu and non-LTR LINE-1 retrotransposition [54]. The observation that this APOBEC3H variant expresses well in mammalian cells compared to wild type was attributed to its highest potential in HIV-1 restriction. In short, the variants of APOBEC3H show different levels of restricting retroviruses and retrotranpositions. Harari *et al.* continued to study the polymorphisms that exist in the human APOBEC3H population and found that more APOBEC3H variants exist in the human population [53].

As explained below, several groups have also provided evidence for the existence of multiple APOBEC3H variants in human tissues and cells. Additionally, the degree of antiretroviral activity against HIV-1 and resistance to vif is assessed in several findings.

Harari *et al.* [53] studied APOBEC3H cDNA obtained from peripheral blood mononuclear cells (PBMC) of twelve HIV-1 negative healthy blood donors. Four different APOBEC3H genetic variants were identified in these donors. HapI-GKE served as the

wild type reference sequence, (NM_181773) where G, K and E are the amino acids at positions at 105, 121 and 178 in human APOBEC3H. The nonsynonymous substitution mutations at amino acid positions at this positions; G105R, K121D and E178D produce the second haplotype of human APOBEC3H, denoted HapII - RDD. Moreover, two additional haplotypes were identified in these cells [53]. Hap II - RDD with Δ15N is denoted as haplotype III. Hap II - RDD with Δ15N and L18R is denoted as haplotype IV. The four major haplotypes are expressed in a population dependent manner. Additional APOBEC3H genetic variants are listed in the Single Nucleotide Polymorphism database at NCBI. (www.ncbi.nlm.nih.gov/projects/SNP). These are one synonymous substitution mutation at codon 43, two deletions (Δ14 and Δ15) and several nonsynonymous mutations (R18L, G37H, G105R, K121E/N, S140G, and E178D). Haplotype I was the most abundant variant of APOBEC3H as it was present in 10 out of 12 blood donors [53].

The alternative splicing of human genes plays an important role in the functions of the gene products leading to diverse and varying enzyme amounts and activities in cells. In addition, the expression profiles and the levels of the splice variants may differ, in a cell or tissue specific manner. Sometimes the expression profile reflects the function of the variant [55]. In addition to the human APOBEC3H haplotypes, four different splice variants were found in the human blood cells, namely sv-154, sv-182, sv-183 and sv-200 [53]. The numbers denote the length of the proteins in amino acids.

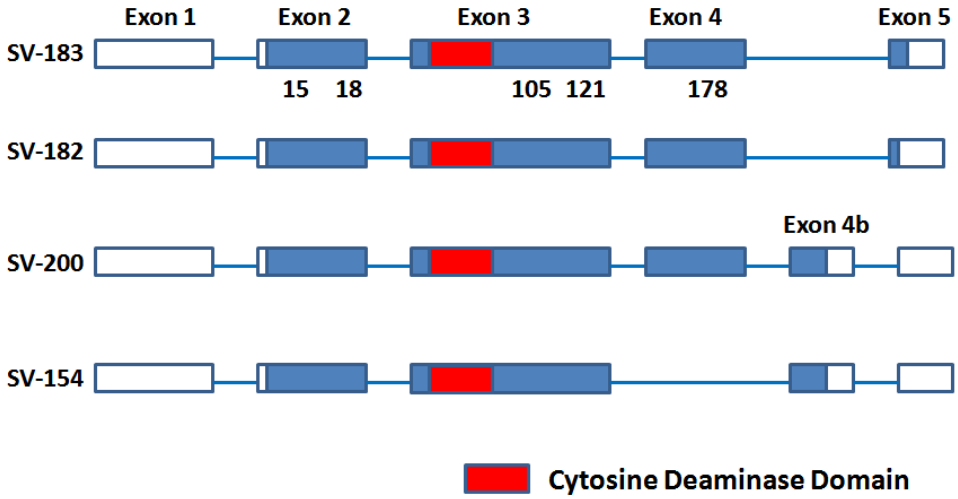

Figure 4: Representation of the four major splice variants of human APOBEC3H. CDD: cytosine deaminase domain.

All the splice variants share exons 2, 3 and 4, except sv-154, which has skipped exon 4 but contains the cryptic exon 4b. In addition to exon 4, sv-200 has the cryptic exon 4b. sv-182 is devoid of the glutamine in exon 5, that is found in sv-183 (Figure 4). The four splice variants are comprised of SNPs that are described above. From the genomic sequences of the transcript variants, it can be concluded that human APOBEC3H is highly diverse and contains multiple haplotypes and splice variants that have distinct C-terminal domains. Harari *et al*. [53] compared the potential of inhibiting HIV-1 by all the variants of human APOBEC3H. They observed that the splice variants of hapII-RDD was more powerful and uniform in restricting HIV-1 than hapI-GKE, as the splice variants of hapI-GKE showed varying levels of antiretroviral ability. But, the sv-183 of hapI-GKE showed 20-fold more restriction potential than background, which was the level of restriction observed by APOBEC3G. This result suggests that sv-183 of HapI-GKE has greater potential in restricting HIV-1 replication in mammalian cells [53].

Gene: APOBEC3H ENSG00000100298

Description	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3H [Source:HGNC Symbol;Acc:24100]
Location	Chromosome 22: 39,493,229-39,500,072 forward strand.
INSDC coordinates	chromosome:GRCh37:CM000684.1:39493229:39500072:1
Transcripts	This gene has 5 transcripts (splice variants) Hide transcript table

Show/hide columns Filter

Name	Transcript ID	Length (bp)	Protein ID	Length (aa)	Biotype	CCDS	GENCODE basic
APOBEC3H-202	ENST00000442487	1047	ENSP00000411754	183	Protein coding	CCDS13985	Y
APOBEC3H-001	ENST00000348946	1046	ENSP00000216123	182	Protein coding	CCDS54531	Y
APOBEC3H-002	ENST00000401756	1023	ENSP00000385741	200	Protein coding	CCDS54530	Y
APOBEC3H-201	ENST00000421988	639	ENSP00000393520	154	Protein coding	CCDS54532	Y
APOBEC3H-003	ENST00000474235	338	No protein product	-	Retained intron	-	-

Figure 5: Human APOBEC3H splice variants from ENSEMBL genome browser

SV-200

MALLTAETFRLQFNNKRRLRRPYYPRKALLCYQLTPQNGST
PTRGYFENKKKCHAEICFINEIKSMGLDETQCYQVTCYLTWS
PCSSCAWELVDFIKAHDHLNLGIFASRLYYHWCKPQQKGLR
LLCGSQVPVEVMGFPKFADCWENFVDHEKPLSFNPYKMLE
ELDKNSRAIKRRLERIK*IPGVRAQGRYMDILCDAEV*

SV-183

MALLTAETFRLQFNNKRRLRRPYYPRKALLCYQLTPQNGST
PTRGYFENKKKCHAEICFINEIKSMGLDETQCYQVTCYLTWS
PCSSCAWELVDFIKAHDHLNLGIFASRLYYHWCKPQQKGLR
LLCGSQVPVEVMGFPKFADCWENFVDHEKPLSFNPYKMLE
ELDKNSRAIKRRLERIK*QS*

SV-182

MALLTAETFRLQFNNKRRLRRPYYPRKALLCYQLTPQNGST
PTRGYFENKKKCHAEICFINEIKSMGLDETQCYQVTCYLTWS
PCSSCAWELVDFIKAHDHLNLGIFASRLYYHWCKPQQKGLR
LLCGSQVPVEVMGFPKFADCWENFVDHEKPLSFNPYKMLE
ELDKNSRAIKRRLERIK*S*

SV-154

MALLTAETFRLQFNNKRRLRRPYYPRKALLCYQLTPQNGST
PTRGYFENKKKCHAEICFINEIKSMGLDETQCYQVTCYLTWS
PCSSCAWELVDFIKAHDHLNLGIFASRLYYHWCKPQQKGLR
LLCGSQVPVEVMGFP*DSRGTCAGSLHGYIV*

Figure 6: Four splice variants of human APOBEC3H found in PBMCs (sv-183, sv-182, sv-200 and sv-154). They contain distinct C-terminal domains The CDD is underlined. In purple, the unique C-terminal domains are indicated.

Although APOBEC3H is distinct from other APOBEC enzymes in humans, it closely resembles the CTD of mouse APOBEC3 [9]. Mouse APOBEC3 is sensitive to HIV-1 vif and the restriction potential of mouse APOBEC3 against HIV-1 is comparable to that of human APOBEC3G [53]. Harari *et al.* [53] and several other groups [56, 57] have confirmed that human APOBEC3H is also HIV-1 vif resistant. To study the underlying mechanism of the resistance to HIV-1 vif by human APOBEC3H, Li *et al.* [57] compared the level of susceptibility of APOBEC3H to several vif proteins cloned from HIV-1 infected patients with different APOBEC3H genotypes. The SNP K121D in hap-II has been identified as the major determinant for conferring resistance against HIV-1 vif, as hapII-RDD has increased antiretroviral activity against HIV-1, consistent with previous studies [53, 57]. Interestingly, change of lysine at position 121 to an acidic amino acid increased susceptibility to HIV-1 vif. Using novel antibodies to detect untagged human APOBEC3H, instead of identifying tagged versions, they observed that the hap-I conferred low resistance to HIV-1, as did splice variants of hap-III and hap-IV [57]. Consistent with previous studies, hap-II was highly expressed than the other variants, and showed increased potential in restricting vif deficient HIV-1. In contrast to the previous studies, by directly comparing protein levels with and without HIV-1 vif, Zhen *et al.* [56] proposed that hap-II is susceptible to HIV-1 vif, although hap-I is completely resistant to proteasomal degradation. Studying a panel of vif mutants, they suggested that, apart from the motifs in vif that are involved in interacting with APOBEC3G and APOBEC3F, additional motifs of vif were identified in interacting with APOBEC3H hap-II. After obtaining similar expression levels for both hap-I and hap-II, Zhen *et al.* [56] observed that both the variants had similar virion packaging levels and antiretroviral activity against HIV-1. As the mutant hap-I G105R was more stable than the hap-I variant, and showed increased steady state

protein levels, it can be concluded that residue 105 is critical for the stability of the protein [56]. Introduction of this positively charged residue at this position might increase its solubility or provide a suitable interaction surface for an accessory factor that increases its stability as this residue is exposed to the surface (From modeling with APOBEC3G-NTD and APOBEC2 structure).

Ooms *et al.* [58] proposed a different phenomenon to explain the varying levels of antiretroviral activity between hap-I and hap-II. The localization of the APOBEC3H variants within the egressing HIV-1 virion was considered as the vital factor, not protein expression levels. Hap-II RDD interacted with the nucleocapsid part of the Gag protein in the virion, in an RNA dependent manner; whereas hap-I GKE interacted with the C-terminal of the capsid. The high affinity of the hap-II variant to host RNA resembles APOBEC3G, as APOBEC3G is incorporated to the virion in an RNA dependent manner. As hap-II RDD is proficient in restricting HIV-1, its ability to interact with the nucleocapsid part of the Gag protein was considered to be the major determinant. In contrast, hap-I did not show any RNA binding or HIV-1 restriction [59]. The W115 residue of hap-II RDD was crucial for the RNA binding ability [59]. The deletion of N15 in hap-III and hap-IV is implicated in its reduced affinity to RNA. Using a rifampicin resistance assay, Ooms *et al.* [58] observed that the APOBEC3H variants have comparable DNA cytosine deaminase activities, but the activity correlated with the expression in *E. coli.*

Wang *et al.* [60] in 2011 identified three novel SNPs in six different human populations, in addition to the N15Δ, R18L, G105R, K121D and E178D, discussed above. The new polymorphisms are named hap-V (N15, R18, R105, D121, E178), hap-VI (N15Δ, L18, G105, K121, D178) and hap-VII (N15, R18, R105, K121, E178). Hap-V and VII expressed high levels of proteins in the steady state and restricted HIV-1 replication. From

the studies on hap-VII, the motif, YYXW (residues 112-115) was identified to be critical for the interaction with the HIV-1 nucleocapsid, hence virion packaging. The Y112A mutation completely abolished the incorporation of APOBEC3H into HIV-1 virions. Replacement of the corresponding region of APOBEC3H with the motif crucial for the vif interaction from APOBEC3G did not increase the vif sensitivity of APOBEC3H [60].

Li *et a.l* [61] discovered that human APOBEC3H hap-II prefers cytoplasmic localization, similar to macaque APOBEC3H, whereas hap-I is primarily retained in the nucleus. The exclusive localization of hap-II in the cytoplasm is attributed to its interaction with host factors that are necessary for HIV-1 restriction. As retroviral reverse transcription occurs in the cytoplasm, the APOBEC3 enzymes must be located there to do this job. Hap-I GKE is translocated to the nucleus by passive diffusion. The putative NES at the C-terminal of the ancestral primate APOBEC3H has been lost due to mutations [61]. Truncation of the C terminal region of the NES in humans and chimpanzee APOBEC3H was observed later in evolution, leading to increased nuclear localization of the APOBEC3H variants, and reducing antiretroviral activity. The single mutation G105R was sufficient to change from nuclear to cytoplasmic localization for hap-I GKE [61]. A summary of antiviral activity, vif sensitivity, and the distribution profiles of the human APOBEC3H haplotypes and its splice variants are shown in Table 1.

APOBEC3H haplotypes/ splice variants	Antiviral activity	Vif sensitivity	Remarks
Ancestral human APOBEC3H (G105R, N15)	Restrict HIV-1, LINE-1, Alu [51]	Sensitive to HIV-1 vif [51]	African populations
hapII-RDD	Inhibit non-LTR LINE-1 [54], Alu[54] retrotransposition, strong HIV-1 restriction [53, 56-58, 62]	All variants resistant to HIV-1 vif [53], completely sensitive to HIV-1 vif [56]	Prevalent in Africans [57, 60]
hap I –GKE sv-183	WT HIV-1 restriction comparable to APOBEC3G [53], weak anti HIV-1 activity [57, 58]	HIV-1 vif resistant [53, 57], sensitive to HIV-2 vif [57]	
hap II-RDD sv-183	WT HIV-1 restriction higher than APOBEC3G [53], powerful anti HIV-1 Δvif activity [57, 61]	HIV-1 vif resistant [53], Somewhat sensitive to HIV-1 vif [57], sensitive to HIV-2 vif [57]	
hap I –GKE sv-182	Poor restriction of HIV-1 than hap I –GKE sv-183 [53]		

APOBEC3H haplotypes/ splice variants	Antiviral activity	Vif sensitivity	Remarks
hap I –GKE sv-200	Poor restriction of HIV-1 than hap I –GKE sv-183 [53]		
hap II –RDD sv-182	Higher restriction of HIV-1 than hap II- RDD sv-183 [53]		
hap II –RDD sv-200	Higher restriction of HIV-1 than hap II- RDD sv-183 (comparable to APOBEC3G) [53]	HIV-1 vif resistant [53]	
hap - III sv-183	No HIV-1 restriction [57]		
hap – IV sv-183	No HIV-1 restriction [57]		
hap – V sv-183	Restrict HIV-1[60]		African-American, Chinese populations[60]
hap – VII sv-183	Restrict HIV-1[60]		European Caucasian[60]

Table 1: Summary of activity and distribution of human APOBEC3H

Despite biological functions such as retroviral restriction and inhibition of retrotransposition, the biochemical properties and the enzyme kinetics of human APOBEC3H are currently unknown. Also a high resolution crystal or NMR structure with or without DNA substrate bound is not yet available for any of the variants of human APOBEC3H. In order to obtain biochemically valid kinetic parameters, homogeneous preparation of the active form of the enzyme is required. The question of structure-activity relationships in various haplotypes can be addressed from a valid crystal or NMR structure. In addition, the residues that are important for interactions with HIV-1 vif in different variants can also be determined by suitable crystal structures with vif. Finally, homogeneous preparations of human APOBEC3H protein variants are crucial for analyzing and confirming the activity profile observed in biological studies.

In the Ensembl Genome Browser (accessed March 2014), four protein coding human APOBEC3H variants are found. The gene identity is ENSG00000100298. For this study, I used two hapI-GKE splice variants that encode 200 and 183 amino acid proteins. The DNA cytosine deaminase activity was compared between the two variants using both an *E. coli* based genetic reversion assay and an *in vitro* DNA cytosine deamination activity assay which was carried out using purified APOBEC3H.

CHAPTER 2: MATERIALS AND METHODS

2.1 Cloning of human APOBEC3H

A cDNA clone of human APOBEC3H hapI-GKE sv-200 (NCBI sequence NM_001166003) was purchased from OriGene Technologies (Rockville, MD) and used as the template for cloning. All the clones that were generated were validated by sequencing. (The DNA Sequencing Core, University of Michigan, Ann Arbor). Plasmid extractions on a small scale were done using an E.Z.N.A.® Plasmid Mini Kit I (Omega Bio-Tek, Norcross, GA).

The cDNA were cloned into the pACYC184-based expression plasmid pSU24 [63] to carry out the kanamycin resistance reversion assay (Figure 7). To clone the splice variant with 200 amino acids, the cDNA was amplified using the primers, pSUA3H200-F and pSUA3H200-R to introduce 5′ *Hind*III and 3’ *EcoR*I sites. To clone, APOBEC3H-183, the same template cDNA was used but was amplified with pSUA3H183-F and pSUA3H183-R and the same restriction sites were introduced. The resulting clones were named as pSUA3H-200 and pSUA3H-183, respectively.

To purify the APOBEC3H-200 protein, an N-terminally GST tagged version was cloned (Figure 8). The cDNA for APOBEC3H hap I GKE sv-200 was cloned into pGEX4T3 vector (GE Healthcare) as 5’ *EcoR*I and 3’ *Xho*I sites and named pGA3H200.

To purify the APOBEC3H-183 protein, an N-terminally poly-histidine tagged version was generated (Figure 9). The cDNA for APOBEC3H hap I GKE sv-200 was used as template and amplified cDNA of sv-183. cDNA was cloned into pET28a+ vector (Novagen) as 5’ *Nde*I and 3’ *Xho*I sites and named pETA3H183.

The primers used for cloning were custom synthesized by Sigma-Aldrich (Woodlands, TX). The sequences of the primers are listed in Table 2.

Name of the oligonucleotide	**Sequence**
pSUA3H200-F	5’ GCG GCG GCG AAG CTT ATG GCT CTG TTA ACA GCC
pSUA3H200-R	5’ GCG GCG GCG GAA TTC TCA TCA GAC CTC AGC ATC
pSUA3H183-F	5’ GCG GCG GCG AAG CTT ATG GCT CTG TTA ACA GCC
pSUA3H183-R	5’ GCG GCG GCG GAA TTC TCA TCA GGA CTG CTT TAT CCT
pETA3H183-F	5’ GGC CGC CCA TAT GGC TCT GTT AAC AG
pETA3H183-R	5’ GCG CGC GGA ATT CTC ATC AGG ACT GCT TTA TCC TCT C
GSTA3H200-F	5’ GGG CCC GGG GAA TTC CGC TCT GTT AAC AGC CG
GSTA3H200-R	5’ GGG CCC GGG CTC GAG TCA GAC CTC AGC ATC ACA
CCC	(6-FAM) 5' ATT ATT ACC CAT TTA TT
CCU	(6-FAM) 5' ATT ATT ACC UAT TTA TT
ACC	(6-FAM) 5' ATT ATT AAC CAT TTA TT
TCC	(6-FAM) 5' ATT ATT ATC CAT TTA TT
GCC	(6-FAM) 5' ATT ATT AGC CAT TTA TT

Table 2: The oligonucleotide sequences.
6-FAM: 6-carboxyfluorescein

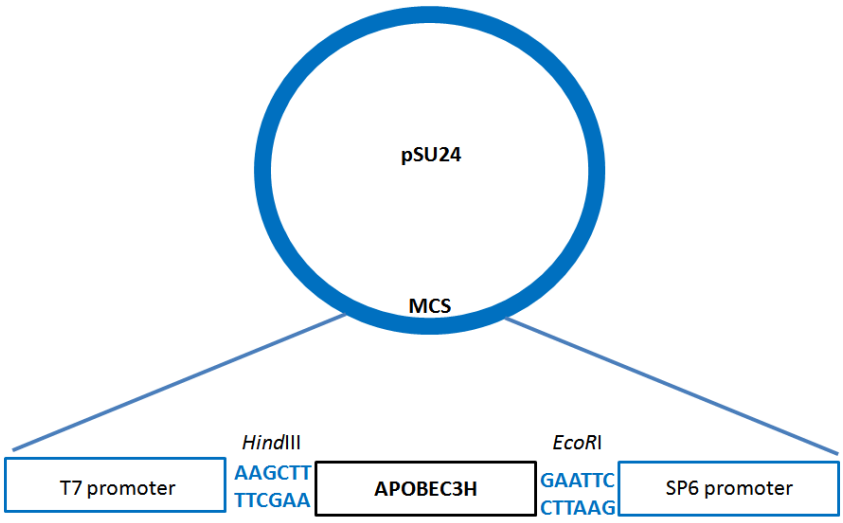

Figure 7: pSU24 plasmid map and MCS.
APOBEC3H splice variants were introduced as *Hind*III and *EcoR*I fragments.

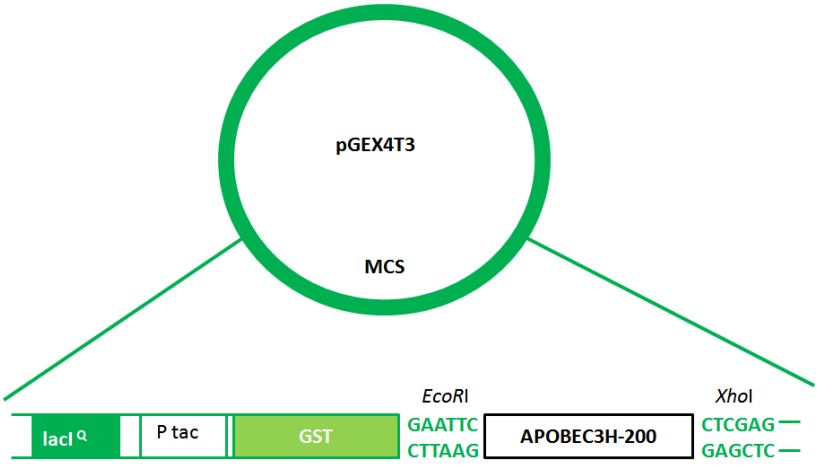

Figure 8: pGEX4T3 plasmid map and MCS.
APOBEC3H-200 was introduced as an *EcoR*I and *Xho*I fragment.

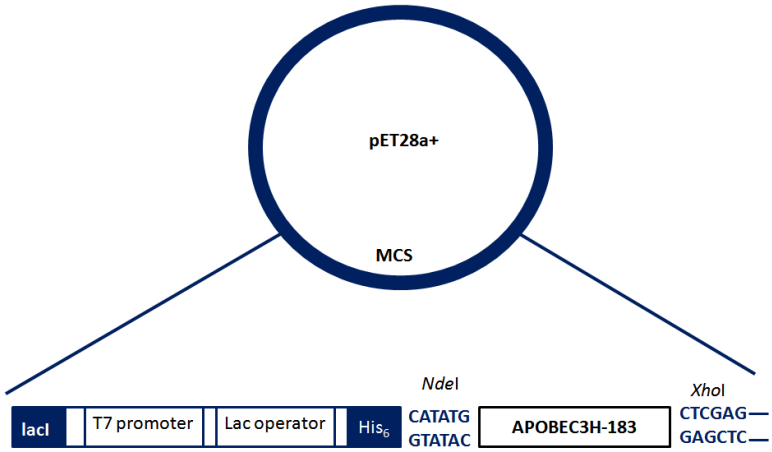

Figure 9: pET28a+ plasmid map and MCS.
APOBEC3H-183 was introduced as an *Nde*I and *Xho*I fragment.

2.2 Purification of GST-APOBEC2H-200

The plasmid pGA3H200 was transformed into the *E. coli* strain BL21. Ten to twelve single colonies were inoculated into a single 5 mL of Luria Bertani (LB) medium culture containing 50 µg/mL carbenicillin, which was grown overnight at 37 °C. A 1:1000 dilution was performed from the overnight culture into 500 mL LB with carbenicillin. The cells expressing the protein were grown to mid-log phase with adequate aeration at 37 °C. The culture was cooled to 4 °C and expression was induced by adding IPTG (Gold Biotechnology) at a final concentration of 0.2 mM. The culture was grown overnight at 16 °C with shaking. Cells were spun down using a Sorvall rotor at 8,000 rpm for 20 minutes at 4 °C. The cells were lysed using a French pressure cell press (Thermo Spectronic) in 30 mL of 20 mM Tris-HCl, pH 8.0 and 50 mM NaCl containing a half a tablet of an EDTA free protease inhibitor cocktail. (Roche diagnostics, Indianapolis. IN). The lysate was clarified by centrifuging at 14,000 rpm for 20 minutes at 4 °C. The cleared lysate was

incubated with glutathione sepharose 4B beads (GE Healthcare) in a total column volume of 750 µL for 2 hours at 4 °C on a rotating platform. The lysate was allowed to settle and the flow through was collected. The beads were washed with 10 column volumes of buffer containing 20 mM Tris-HCl, pH 8.0 and 50 mM NaCl. The protein was eluted using buffer containing 50 mM Tris-HCl, pH 8.0 and 20 mM reduced glutathione (Sigma Aldrich); 1 mL fractions were collected. The cell free lysate (15 µL each), flow through, wash fractions and elution fractions were mixed with SDS-PAGE gel loading buffer (50 mM Tris-HCl, pH 6.8, 100 mM DTT, 2% SDS, 0.1% bromophenol blue and 10 % glycerol). The samples were heated at 95 °C for 5 minutes and briefly spun down. The soluble portion was loaded onto a 12% SDS-PAGE gel and ran at 100 V in 20mM Tris-HCl, 192 mM glycine and 0.1% SDS buffer. The gel was stained in Coomassie Brilliant Blue (BioRad) (1g/L), 50% methanol and 10% glacial acetic acid. The gel was destained using 50% methanol and 10% glacial acetic acid. The fractions that contained the desired protein (expected size of GST-A3H200 is ~48 kDa) were dialyzed using 10 kDa molecular weight cut-off Slide-A-Lyzer Dialysis Cassettes (Thermoscientific, Rockford, IL) in 20 mM Tris-HCl, pH 7.6, 50 mM NaCl, 1 mM EDTA, 1 mM DTT and 10% glycerol at 4 °C overnight. The resultant protein preparation was concentrated using a 10 kDa molecular weight cut-off Amicon Ultra Centrifugal device (Milipore, Billerica, MA) by centrifuging at 4,000 g for 20-30 minutes at 4 °C. The concentration of the final protein preparation was measured using Bradford reagent (Bio-Rad) against BSA standards from the absorption at 595 nm. Protein was aliquoted and stored at -20 °C until further use. For initial activity assays, 1 µg of the protein was used in each reaction.

2.3 Western Blotting of GSTA3H200

To confirm the overexpression of GSTA3H200 in BL21 cells, a Western blot was carried out on the concentrated eluate obtained from the glutathione affinity column. Briefly, 20 µL of the elution fraction was mixed with of SDS-PAGE gel loading buffer (50 mM Tris-HCl, pH 6.8, 100 mM DTT, 2% SDS, 0.1% bromophenol blue and 10 % glycerol). The content was mixed and heated at 95 °C for 5 minutes and briefly spun down. An aliquot of the soluble portion (15 µL) was loaded onto a 12% SDS-PAGE gel and ran at 100V in 20 mM Tris-HCl, 192 mM glycine and 0.1% SDS.

The proteins were electrophoretically transferred to a PVDF membrane (EMD Millipore) which had been presoaked in methanol. The transfer was carried out at 100 mA for two hours at 4 °C. The blot was incubated with the blocking buffer, 5% non-fat dry milk in 1X TBS-T (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Tween-20) at 4 °C overnight. The blot was washed for 5 minutes three times with TBS-T and incubated with the rabbit anti-GST tag primary antibody (GeneScript) with a final concentration of 1 µg/mL at room temperature for two hours. The blot was washed again with TBS-T and was incubated with the secondary antibody, goat anti-rabbit antibody (AnaSpec) with a final concentration of 1 µg/mL at room temperature for one hour. The blot was washed again with TBS-T and scanned using Typhoon Phosphorimager (GE Healthcare) at the excitation wavelength of 633 nm under a red laser.

2.4 Kanamycin resistance reversion assay

The *E. coli* based genetic assay was described before [64, 65]. Briefly, a defective *kanamycin* allele was used; in which 94th leucine residue was mutated to proline (kanamycin sensitive phenotype in *E. coli*). After one round of replication, the deamination of cytosines to uracils in this codon will result in C to T mutations, restoring the kanamycin resistant phenotype. The number of C to T mutations is scored as the kanamycin resistance revertant frequency (Figure 10). Plasmids, pSUA3H-200 or pSUA3H-183 was cotransformed with the pUP31 plasmid into BH214 strain of *E. coli* which is a GM31 *ung mug* strain. The defective *kanamycin* allele is provided from pUP31 [66, 67]. For the positive control, human APOBEC3A cloned into pSU24 (pSUA3A) was used [26]. The empty vector, pSU24, or the catalytically null mutant of human APOBEC3A cloned in pSU24 (pSUA3AE72A) [26] were used as negative controls. The cotransformants were selected by plating on an LB plate containing chloramphenicol (25 µg/mL) and carbenicillin (50 µg/mL). LB medium (5 mL) with chloramphenicol (25 µg/mL) and carbenicillin (50 µg/mL) was inoculated with a single colony from the selective plate. The culture was grown overnight. A 1:1000 dilution of the overnight culture was carried out in LB containing the appropriate antibiotics and the culture was grown for two hours at 37 °C with shaking. For the experiments with IPTG induction, varying concentrations of IPTG were added to the cultures. The cultures were split into eight or more independent cultures (5 mL each). The cultures were grown until the mid-log phase (i.e. OD_{600} ~ 0.4 to 0.6). The cultures were spread onto LB-agar plates with 50 µg/mL kanamycin to score the number of revertants. To score for the total number of viable cells, a 2.5 x 10^6 dilution of the cultures were plated on LB-agar plates containing 25 µg/mL carbenicillin. Kanamycin-resistance revertant frequency was calculated as the ratio between the numbers of

kanamycin-resistant colonies to the total number of viable cells. The median values were calculated and analyzed using GraphPad Prism software. The Mann-Whitney t-test was used to calculate p-values.

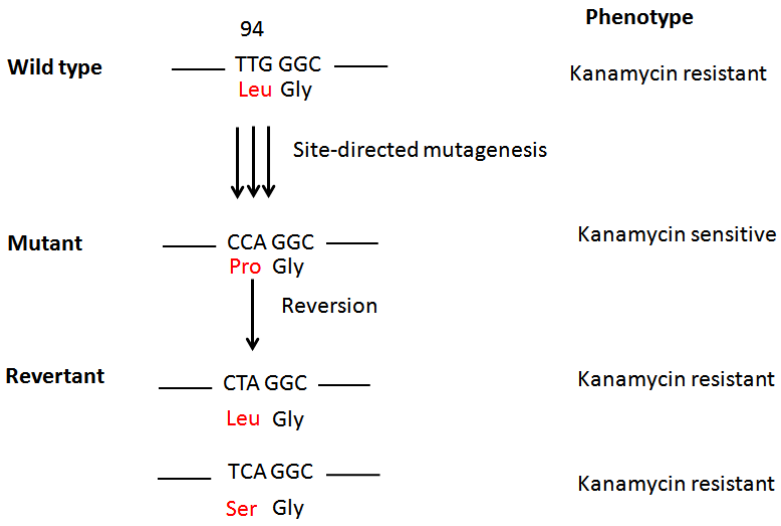

Figure 10: The basis of the kanamycin resistance reversion assay.

2.5 Purification of poly-histidine tagged A3H-183 using nickel affinity

The plasmid pETA3H183 was transformed into the *E. coli* strain BL21 (DE3). Ten to twelve single colonies were inoculated into a single 5 mL of LB medium culture containing 50 µg/mL kanamycin, which was grown overnight at 37 °C. A 1:1000 dilution was performed from the overnight culture into 500 mL LB with kanamycin. The cells expressing the protein were grown to mid-log phase with adequate aeration at 37 °C. The culture was cooled to 4 °C and expression was induced by adding IPTG (Gold Biotechnology) at a final concentration of 0.2 mM. The culture was grown overnight at 16 °C with shaking. Cells were spun down using a Sorvall rotor at 8,000 rpm for 20 minutes at 4 °C. The cells were lysed using French pressure cell press (Thermo Spectronic) in 30 mL of 20mM Tris-HCl pH 8.0, 50mM NaCl and 40mM imidazole (Sigma-Aldrich) containing a half a tablet of the EDTA free protease inhibitor cocktail. (Roche diagnostics,

Indianapolis. IN). The lysate was clarified by centrifuging at 14,000 rpm for 20 minutes. The cleared lysate was incubated with high affinity Ni-charged resin (GenScript) in a total column volume of 1 mL for one hour at 4 °C on a rotating platform. The lysate was allowed to settle and the flow through was collected. The beads were washed with 10-15 column volumes of the buffer containing 20 mM Tris-HCl, pH 8.0, 40 mM imidazole and 50 mM NaCl (wash 1). Wash was repeated with 20 mM Tris-HCl, pH 8.0, 40 mM imidazole with increasing NaCl concentrations (200 mM and 500 mM respectively) with each step. The wash steps were carried out until no detectable proteins were eluted in Bradford qualitative assay. These wash steps are indicated as wash 2 and 3 in Figure 16. The protein was eluted in 20 mM Tris-HCl, pH 8.0, 500 mM NaCl with 300 mM imidazole. 15 µL of the cell free lysate, flow through, wash and elution fractions were mixed with SDS-PAGE gel loading buffer. The samples were heated at 95 °C for 5 minutes and briefly spun down. The soluble portion was loaded onto a 15% SDS-PAGE gel and ran at 100V in 20mM Tris-HCl, 192 mM glycine and 0.1% SDS buffer. The gel was stained in Coomassie Brilliant Blue (BioRad) (1g/L), 50% methanol and 10% glacial acetic acid. The gel was destained using 50% methanol and 10% glacial acetic acid. The fractions that contained the desired protein (expected size of His-A3H183 is ~23 kDa) were dialyzed using 10 kDa molecular weight cut-off Slide-A-Lyzer Dialysis Cassettes (Thermoscientific, Rockford, IL) in a step-wise manner. First, the dialysis was carried out in 20 mM Tris-HCl, pH 7.6, 200 mM NaCl, 1mM EDTA, 1mM DTT and 10% glycerol at 4 °C for 8-10 hours and it was repeated for another 8 hours, with the same buffer, except with a NaCl concentration of 50 mM. The resultant protein preparation was concentrated using 10 kDa molecular weight cut-off Amicon Ultra Centrifugal device (Milipore, Billerica, MA) by centrifuging at 4000 g for 20-30 minutes at 4 °C. The concentration of the final protein

preparation was measured using Bradford reagent against a standard series of BSA from the absorption at 595 nm. The protein was aliquoted and stored at -20 °C until further use. For initial activity assays, 1 µg of the protein was used in each reaction.

2.6 Western Blotting of his-A3H183

To confirm the overexpression of N-terminally poly-histidine tagged APOBEC3H-183 in BL21(DE3) cells, a Western blot was carried out with the eluate and other fractions obtained from nickel affinity purification. Briefly, 17µL of the cell free lysate, flow-through, portions from the wash steps, and elution fractions were each mixed with SDS-PAGE gel loading buffer. The samples were mixed and heated at 95 °C for 5 minutes and briefly spun down. Aliquots of the soluble portion (15 µL) were loaded onto a 15% SDS-PAGE gel and ran at 100V in 20 mM Tris-HCl, 192 mM glycine and 0.1% SDS.

The proteins were electrophoretically transferred the proteins to a PVDF membrane (EMD Millipore) which has been presoaked in methanol. The transfer was carried out at 100 mA for two hours at 4 °C. The blot was incubated with the blocking buffer, 1% BSA in 1X TBS-T (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Tween-20) at room temperature overnight. The blot was washed for 5 minutes three times with TBS-T and incubated with the mouse anti-his tag primary antibody (AnaSpec) with a final concentration of 1 µg/mL at room temperature for two hours. The blot was washed again with TBS-T and was incubated with the secondary antibody, goat anti-mouse IgG (H+L), Hilyte Flour^{TM} 647-labeled (AnaSpec) with a final concentration of 1 µg/mL at room temperature for one hour. The blot was washed again with TBS-T and scanned using a Typhoon Phosphorimager (GE Healthcare) at the excitation wavelength of 633 nm under a red laser.

2.7 Size exclusion chromatography of poly-histidine tagged A3H-183

A fresh preparation of overexpressed lysate was obtained from a 2 L culture, as described above in section 2.5. All growth and purification conditions were scaled-up accordingly. As the protein was active under the elution condition that was used in nickel affinity purification, the buffer containing 20 mM Tris-HCl, pH 8.0 and 500 mM NaCl was used to equilibrate the column. A total of ~0.8 mg of crude eluate from the nickel affinity column was used to load a HiLoad 16/600 Superdex 200 pg (GE Healthcare) size exclusion column. Fractions (2 mL) were collected at a flow rate of 0.5 mL per minute. The fractions containing the refined protein were pooled and a total of 10 µL was loaded onto a 15% SDS-PAGE gel. Fractions containing the higher molecular weight contaminant were also processed as above and 10 µL was loaded onto the gel. After confirming the presence of the desired protein at the proper void volume, without the high molecular weight contaminant, the pooled fractions were concentrated using a 10 kDa molecular weight cut-off Amicon Ultra Centrifugal device, (Milipore, Billerica, MA) by centrifuging at 4,000 g for 20-30 minutes at 4 °C. The two step dialysis was carried out as described in section 2.5. The high molecular weight contaminant was removed from the size exclusion column. The resultant pure protein was used to carry out further activity assays.

2.8 *In vitro* activity assays of purified APOBEC3H proteins

To complement the genetic data obtained from the kanamycin resistance reversion assay, the partially purified GSTA3H200 and poly-histidine tagged A3H-183 (HisA3H183) was tested for their ability to deaminate cytosine *in vitro.* Essentially pure hisA3H183 obtained from the size exclusion column was also tested. Four different 17-mers of DNA were used in the assay. The oligonucleotides were 5’ 6-FAM labeled and contains two

consecutive cytosines preceded by either A, G, T or C. The oligonucleotide used for a positive control contained 'CCU' sequence. All the oligonucletides were custom synthesized from Integrated DNA Technologies (San Jose, CA) and were PAGE purified prior to use. The sequences of the oligonucleotides are listed in Table 2. Briefly, 1 pmol of oligonucleotide was incubated with partially purified GSTA3H200 or hisA3H183 and incubated for two hours at 37 °C. The total reaction volume was 10 µL in the reaction buffer containing 40 mM Tris-HCl, pH 7.4, 5 mM EDTA, 1 mM DTT and 40 mM NaCl. One unit of *E. coli* UDG (New England Biolabs, Ipswich, MA) was added with the UDG reaction buffer (20 mM Tris-HCl, pH 8, 1 mM EDTA, 1 mM DTT) and the reaction was continued at 37 °C for one hour. NaOH was added to a final concentration of 0.1 M and the reaction was heated for 5 minutes at 95 °C. An equal volume of formamide containing gel loading dye (95% formamide, 0.025% bromophenol blue, 5 mM EDTA) was added to the reaction mixture and the mixture was heated again for 5 minutes at 95 °C. The reaction was briefly spun down at 10,000 g for one minute and 10 µL was loaded onto a denaturing 20% DNA-PAGE gel. The gel was run in 1X TBE buffer at 90 V. The gel was scanned using Typhoon 9210 phosphorimager (GE Healthcare) with an excitation wavelength of 499 nm. The intensities of the substrates and the products were quantified using ImageQuant software. Percent conversion of the substrate was calculated per the equation: (intensity of the product band - intensity of the background) / ((intensity of the substrate band - intensity of the background) + (intensity of the product band - intensity of the background)). For the homogenous protein preparation of hisA3H183 obtained from size exclusion, different amounts were used for the activity assays. The assays were carried out as described above.

CHAPTER 3: RESULTS AND DISCUSSION

The main objective of this project was to obtain a version of human APOBEC3H that is active, both *in vivo* and *in vitro.* To study the DNA cytosine deaminase activity, I used a novel *E. coli* mutator assay, the kanamycin resistance reversion assay.

The kanamycin resistance reversion assay data for the 200 amino acid version of APOBEC3H (Figure 11) suggests that the APOBEC3H-200 is not an efficient DNA cytosine deaminase. Eight to twelve independent cultures were used to determine the median value of the kanamycin resistance reversion frequency of each population, including the negative control (empty vector). The median value of the kanamycin resistance reversion frequency of the splice variant APOBEC3H-200 was only ~3 fold (p=0.052) higher than the empty vector control.

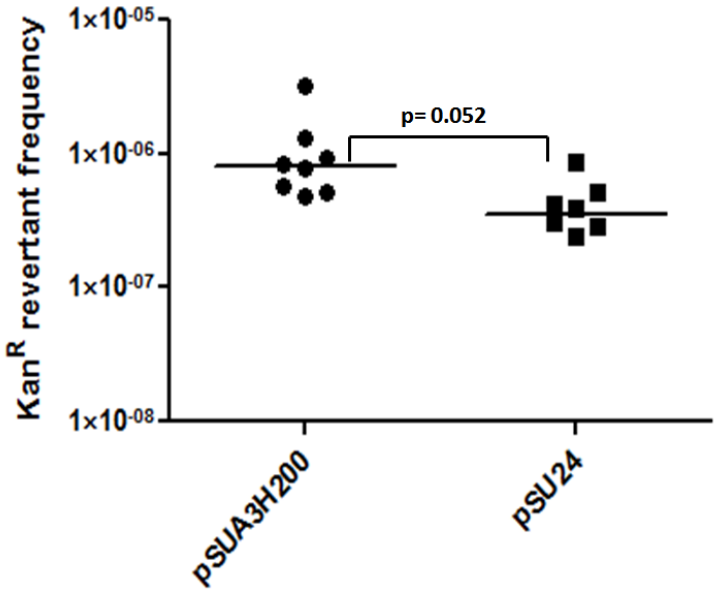

Figure 11: Kanamycin resistance reversion assay results for pSUA3H-200.
The median value of the revertant frequency is shown by the horizontal bar.

To assess the *in vitro* biochemical activity, an N-terminally GST tagged version of A3H200 was obtained using glutathione affinity chromatography. Figure 12 represents the 12% SDS PAGE gel with the protein fractions that were collected.

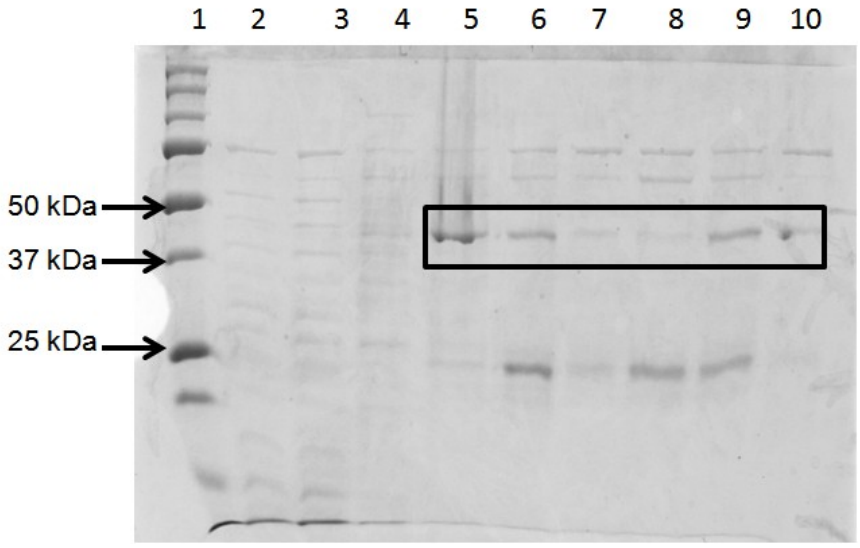

Figure 12: 12% SDS-PAGE gel for GSTA3H200 purification.
Lane 1 : Molecular weight marker, Lane 2 : CFL, Lane 3 : Flow-through,
Lane 4 : wash, Lanes 5-10 : eluates of GSTA3H200 (boxed)

The expected protein size of GSTA3H200 was ~48 kDa, which was not significantly over-expressed. Thus, a Western blot was carried out using anti-GST mouse primary antibody. As presented in Figure 12, the protein seems to be expressed, with smaller molecular weight degradation products. The activity assay was performed as described above using concentrated eluates obtained from the affinity column. A total of 1.5 µg of protein did not give any cleaved DNA product (Figure 14).

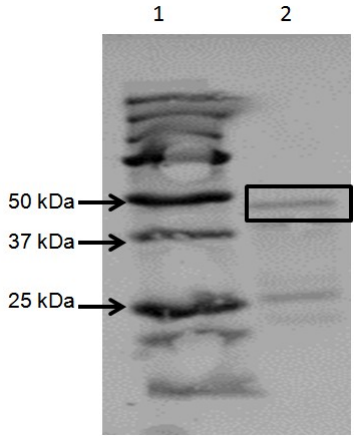

Figure 13: Western blot for GSTA3H200
Lane 1: Molecular weight marker, Lane 2: Eluate from GST affinity column.

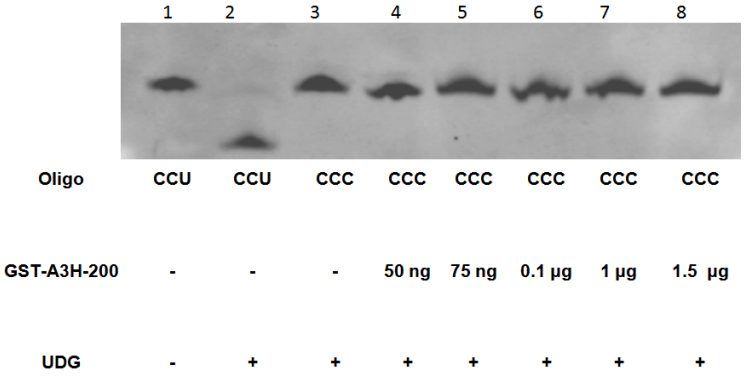

Figure 14: DNA cytosine deaminase activity assay for partially purified GSTA3H200

As human APOBEC3H is a known DNA cytosine deaminase, which contains a conserved zinc binding deaminase domain, we wanted to study the other splice variants of this enzyme. From a previous study by Harari *et al.* [53], the splice variant that encodes 183 amino acids (hapI-GKE-183) displayed ~20 fold higher HIV-1 restriction potential than the other variants. To study the DNA cytosine deaminase efficiency of this splice variant, the kanamycin resistance reversion assay was carried out using the same conditions as

above. Interestingly, cells with the human APOBEC3H-183 showed higher kanamycin resistance revertant frequency than the empty vector control, although the value was not as high as the human APOBEC3A (Figure 15). Human APOBEC3A cloned in pSU24 has shown increased DNA cytosine deaminase activity in this assay [26]. The negative control was a catalytic mutant. APOBEC3AE72A. With increasing concentrations of IPTG, the revertant frequency was also higher. It is likely that higher IPTG concentrations result in greater expression of APOBEC3H and hence higher deamination activity. Addition of IPTG to a final concentration of 100 μM increased the kanamycin resistance reversion frequency ~10 fold.

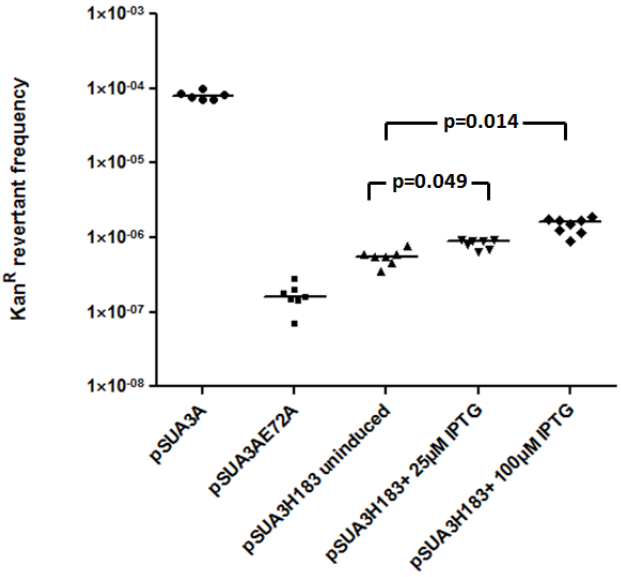

Figure 15: Kanamycin resistance reversion assay results for pSUA3H183.
The median value is shown by the horizontal bar.

To complement the genetic data obtained from the *E. coli* mutator assay, the activity of the APOBEC3H-183 was tested *in vitro* using purified proteins. As the first step in carrying out this *in vitro* activity assay, APOBEC3H-183 was purified as an N-terminally poly-histidine tagged APOBEC3H-183 (His-A3H183) using nickel affinity chromatography.

The SDS-PAGE gel is shown in Figure 16. The crude preparation of protein obtained from a 500 mL of culture was used to carry out the activity assay.

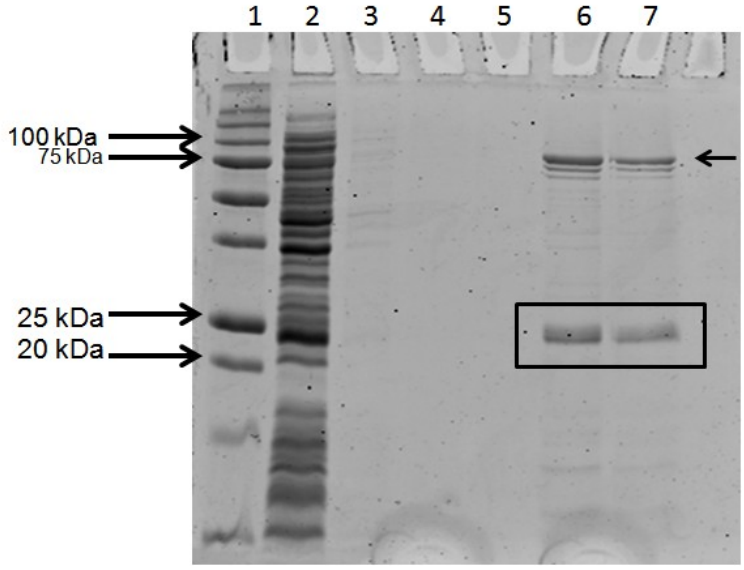

Figure 16: 15% SDS-PAGE gel for nickel affinity purified APOBEC3H-183.
Lane 1: Molecular weight marker, Lane 2: Flow-through, Lane 3: wash 1, Lane 4: wash 2, Lane 5: wash 3, Lane 6-7: eluate. The expected size of his-A3H-183 (~23 kDa) is boxed. A higher molecular weight contaminant is indicated by an arrow to the right.

To confirm the overexpression of the poly-histidine tagged A3H183, a Western blot was carried out using mouse anti-his tag primary antibody. As shown in Figure 17 (lane 9), the correct size of the protein was detected (~23 kDa). This result confirmed that sufficient protein expression in the soluble fraction of the lysate occurred. Unfortunately, a significant amount of the protein was lost and cleared through the wash and flow-through steps, perhaps due to insufficient binding to the nickel beads. This step needs further optimization.

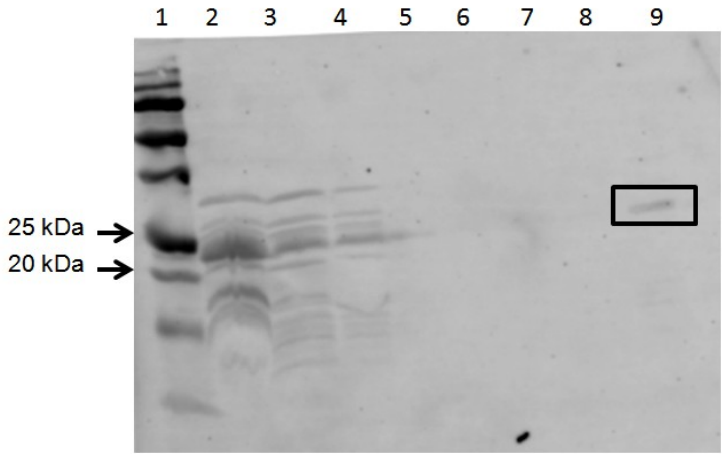

Figure 17: Western blot for protein fractions obtained in nickel affinity purification.
Lane 1: Molecular weight marker, Lane 2: Cell-free lysate, Lane 3: flow-though, Lane 4: wash 1, Lane 5: wash 2, Lane 6: wash 3, Lane 7-8: empty, Lane 9: eluate. The expected protein size of his-APOBEC3H-183 (~23 kDa) is boxed.

We next carried out the *in vitro* biochemical activity assay using the partially purified APOBEC3H-183, described in [68] and [69]. Briefly, an event of DNA cytosine deamination in a single-stranded oligonucleotide will generate uracil. Incubation the olionucleotide with UDG, excises uracil and generates an abasic site. Treatment with 0.1 M NaOH and heating to 95 °C (hot alkaline treatment), cleaves the phosphodiester backbone at the abasic site and introduces a single strand break, resulting in a shorter product. The appearance of this shorter product indicates a positive result or a proficient DNA cytosine deaminase activity of the enzyme tested. The corresponding 20% DNA-PAGE gels are shown in Figures 18 and 19, for assays carried out using the partially purified N-terminal poly-histidine tagged APOBEC3H. The different reactions in each lane are indicated. The appearance of two product bands is attributed to the two deamination events in the consecutive cytosines in the substrate. As seen clearly, the appearance of the shorter product and the complete or partial disappearance of the substrate, corresponding to the amount of enzyme, indicates that the enzyme is active under the

conditions used. The lanes corresponding to the positive and negative controls are indicated. Although the yield was low, the protein preparation was active and was able to deaminate the cytosines in the oligonucleotide. The ‘CCC’ and ‘TCC’ substrates (1 pmol each) were converted 100% to product by 1 µg of the protein. Only ~35% of the substrate was converted in the ‘ACC’ oligomer. Approximately 65% of the ‘GCC’ substrate was converted to product under the same conditions. By averaging the four conversions, the specific activity of the enzyme was calculated to be >0.0125 pmol/µg/min under the specified conditions, suggesting that the enzyme is active, but has a slow turnover. For the ‘CCC’ oligo, the specific activity was >0.017 pmol/µg/min.

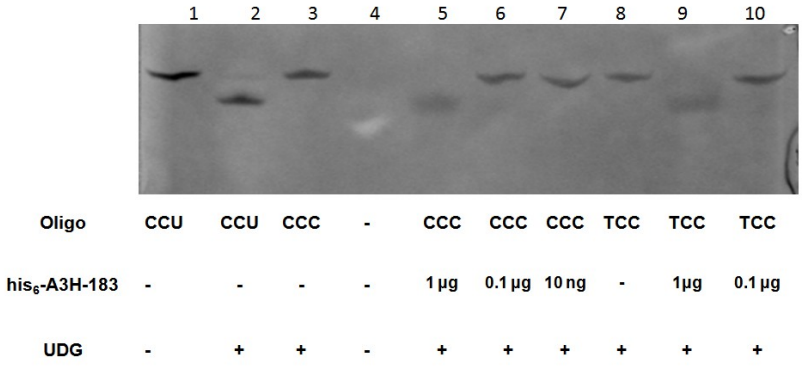

Figure 18: Biochemical activity assay for partially purified APOBEC3H-183.
‘CCC’ and ‘TCC’ oligomers were used. Lane 1 and lane 2 represent the negative and positive controls, respectively. The sequences of the substrate oligonucleotides are indicated in Table 2.

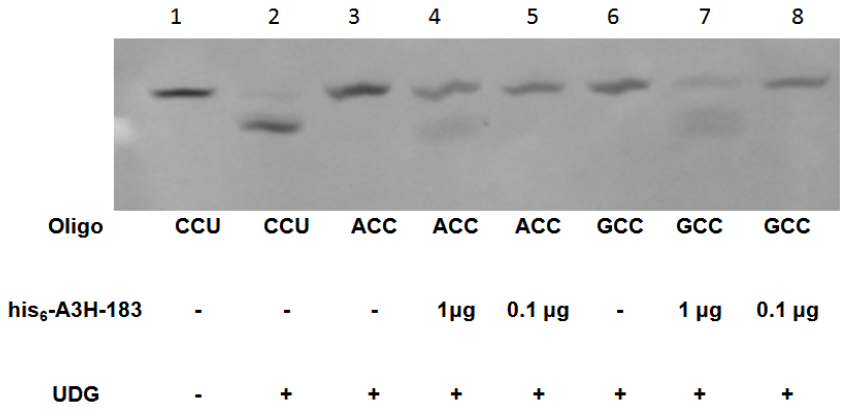

Figure 19: Biochemical activity assay for partially purified APOBEC3H-183.
'ACC' and 'GCC' oligomers were used. Lane 1 and lane 2 represent the negative and positive controls, respectively. The sequences of the substrate oligonucleotides are indicated in Table 2.

In order to further characterize the structure and the catalytic mechanism of the human APOBEC3H hap I GKE sv-183, essentially homogeneous preparation of the active enzyme is required. As the first step, the partially purified protein (his_6-A3H-183) was further refined using size exclusion chromatography. The fractions obtained from the size exclusion column were concentrated using centrifugal devices and analyzed on a 15% SDS-PAGE gel. As shown in the chromatogram (Figure 20), the ~100 kDa contaminant was removed by the size exclusion column. Nearly pure protein was obtained (Figure 21), even though the yield was poor. Varying amounts of the protein were used to carry out the *in vitro* activity assay (Figure 22).

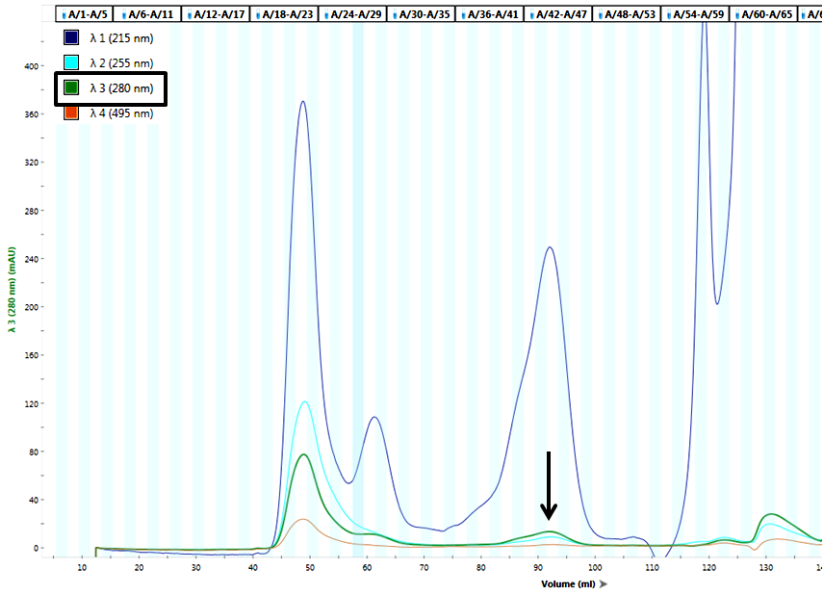

Figure 20: Size exclusion chromatogram.
The peak corresponding for ~23 kDa is indicated by the arrow.

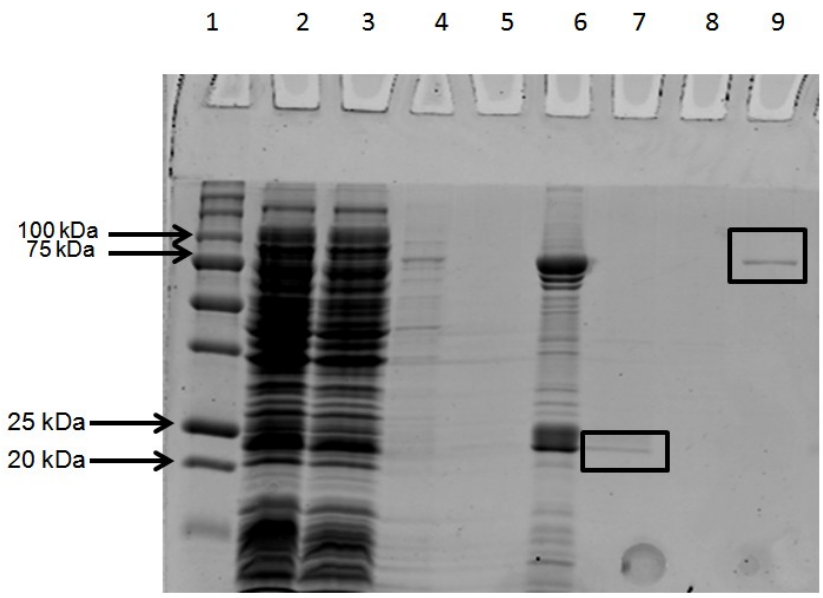

Figure 21: SDS-PAGE gel (15%) with APOBEC3H-183 from a size exclusion column. Lane 1: Molecular weight marker, Lane 2: Cell free lysate, Lane 3: Flow-through, Lane 4: wash 1, Lane 5: wash 2, Lane 6: Eluate from nickel affinity column, Lane 7: concentrated fractions corresponding to ~23 kDa from the size exclusion column, Lane 9: Higher molecular weight contaminant from size exclusion column.

Figure 22 shows the 20% DNA-PAGE gel from the cytosine deaminase activity assay that was carried out using the fully purified his_6-APOBECH-183. Increasing amounts of the enzyme were used. Only 45 ng of enzyme was able to deaminate ~70% of the substrate. For 100% activity, >0.1 µg of the protein was sufficient. The amount of enzyme in the reaction directly correlated to the activity in a 2 hour incubation. The specific activity of the enzyme is increased to ~0.13 pmol/µg/min.

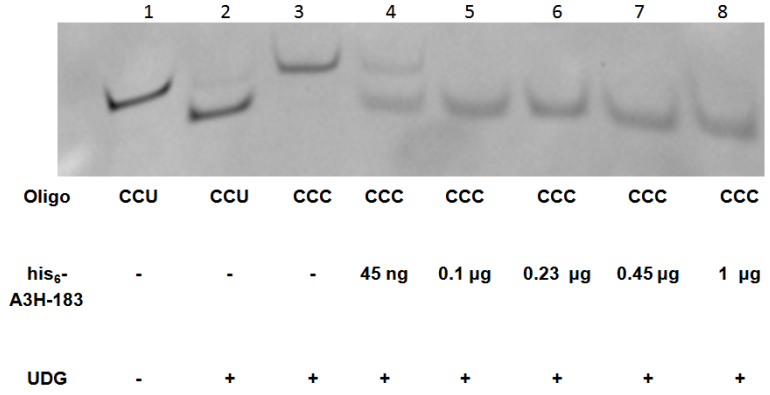

Figure 22: DNA cytosine deaminase activity assay of his-APOBEC3H-183. Different amounts of the protein were used in the assay, as indicated.

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS

To obtain biochemically valid conclusions and to study structure-activity relationship, isolation of catalytically active proteins is necessary. Human APOBEC3H is a good example to study the evolution and positive selection of proteins. As the APOBEC3H shows human population-specific distribution and several splice variants, further biochemical characterization of multiple variants are required. In addition, identification of other cellular partners that interact with this enzyme is crucial for functional analysis. High resolution crystal and NMR structures of human APOBEC3H have not been obtained. I have cloned human APOBEC3H hapI-GKE splice variant 183 as an N-terminally poly-histidine tagged version into a pET28 vector and purified it using nickel affinity and size exclusion chromatography, as the first step in characterizing the biochemical properties of this enzyme. Purified enzyme was tested for its DNA cytosine deaminase activity. The protein was efficient in deaminating cytosine *in vitro*. Further work is required to obtain the enzymatic parameters, such as catalytic efficiency. In addition, the calatytically-null mutant of human APOBEC3H, E56A [70] needs to be isolated and tested *in vitro.* The work in this project needs to be extended to obtain human APOBEC3H-183 in higher yields to characterize it biochemically and structurally. The catalytic mechanisms of the variants can also be evaluated from the structures obtained with suitable DNA substrates.

REFERENCES

1. Chester, A., *et al*., *RNA editing: cytidine to uridine conversion in apolipoprotein B mRNA.* Biochim Biophys Acta, 2000. **1494**(1-2): p. 1-13.
2. Petersen-Mahrt, S.K. and M.S. Neuberger, *In vitro deamination of cytosine to uracil in single-stranded DNA by apolipoprotein B editing complex catalytic subunit 1 (APOBEC1).* J Biol Chem, 2003. **278**(22): p. 19583-6.
3. Sohail, A., *et al*., *Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations.* Nucleic Acids Res, 2003. **31**(12): p. 2990-4.
4. Muramatsu, M., *et al*., *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, 1999. **274**(26): p. 18470-6.
5. Dickerson, S.K., *et al*., *AID mediates hypermutation by deaminating single stranded DNA.* J Exp Med, 2003. **197**(10): p. 1291-6.
6. Muramatsu, M., *et al*., *Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme.* Cell, 2000. **102**(5): p. 553-63.
7. Samaranayake, M., *et al*., *Evaluation of molecular models for the affinity maturation of antibodies: roles of cytosine deamination by AID and DNA repair.* Chem Rev, 2006. **106**(2): p. 700-19.
8. LaRue, R.S., *et al*., *The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals.* BMC Mol Biol, 2008. **9**: p. 104.

9. Conticello, S.G., *et al*., *Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases.* Mol Biol Evol, 2005. **22**(2): p. 367-77.

10. Harris, R.S. and M.T. Liddament, *Retroviral restriction by APOBEC proteins.* Nat Rev Immunol, 2004. **4**(11): p. 868-77.

11. Bishop, K.N., *et al*., *Cytidine deamination of retroviral DNA by diverse APOBEC proteins.* Curr Biol, 2004. **14**(15): p. 1392-6.

12. Wiegand, H.L., *et al*., *A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins.* EMBO J, 2004. **23**(12): p. 2451-8.

13. Wedekind, J.E., *et al*., *Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business.* Trends Genet, 2003. **19**(4): p. 207-16.

14. Browne, E.P. and D.R. Littman, *Species-specific restriction of apobec3-mediated hypermutation.* J Virol, 2008. **82**(3): p. 1305-13.

15. Rulli, S.J., Jr., *et al*., *Interactions of murine APOBEC3 and human APOBEC3G with murine leukemia viruses.* J Virol, 2008. **82**(13): p. 6566-75.

16. Nair, S., *et al*., *Biochemical and biological studies of mouse APOBEC3.* J Virol, 2014. **88**(7): p. 3850-60.

17. Sanchez-Martinez, S., *et al*., *Studies on the restriction of murine leukemia viruses by mouse APOBEC3.* PLoS One, 2012. **7**(5): p. e38190.

18. Mikl, M.C., *et al*., *Mice deficient in APOBEC2 and APOBEC3.* Mol Cell Biol, 2005. **25**(16): p. 7270-7.

19. Koning, F.A., *et al*., *Target cell-mediated editing of HIV-1 cDNA by APOBEC3 proteins in human macrophages.* J Virol, 2011. **85**(24): p. 13448-52.

20. Peng, G., *et al*., *Myeloid differentiation and susceptibility to HIV-1 are linked to APOBEC3 expression.* Blood, 2007. **110**(1): p. 393-400.

21. Vartanian, J.P., *et al*., *Evidence for editing of human papillomavirus DNA by APOBEC3 in benign and precancerous lesions.* Science, 2008. **320**(5873): p. 230-3.

22. Narvaiza, I., *et al*., *Deaminase-independent inhibition of parvoviruses by the APOBEC3A cytidine deaminase.* PLoS Pathog, 2009. **5**(5): p. e1000439.

23. Landry, S., *et al*., *APOBEC3A can activate the DNA damage response and cause cell-cycle arrest.* EMBO Rep, 2011. **12**(5): p. 444-50.

24. Land, A.M., *et al*., *Endogenous APOBEC3A DNA cytosine deaminase is cytoplasmic and nongenotoxic.* J Biol Chem, 2013. **288**(24): p. 17253-60.

25. Stenglein, M.D., *et al*., *APOBEC3 proteins mediate the clearance of foreign DNA from human cells.* Nat Struct Mol Biol, 2010. **17**(2): p. 222-9.

26. Wijesinghe, P. and A.S. Bhagwat, *Efficient deamination of 5-methylcytosines in DNA by human APOBEC3A, but not by AID or APOBEC3G.* Nucleic Acids Res, 2012. **40**(18): p. 9206-17.

27. Suspene, R., *et al*., *Efficient deamination of 5-methylcytidine and 5-substituted cytidine residues in DNA by human APOBEC3A cytidine deaminase.* PLoS One, 2013. **8**(6): p. e63461.

28. Burns, M.B., N.A. Temiz, and R.S. Harris, *Evidence for APOBEC3B mutagenesis in multiple human cancers.* Nat Genet, 2013. **45**(9): p. 977-83.

29. Burns, M.B., *et al*., *APOBEC3B is an enzymatic source of mutation in breast cancer.* Nature, 2013. **494**(7437): p. 366-70.

30. Li, D., *et al*., *Core-APOBEC3C chimerical protein inhibits hepatitis B virus replication.* J Biochem, 2011. **150**(4): p. 371-4.

31. Kock, J. and H.E. Blum, *Hypermutation of hepatitis B virus genomes by APOBEC3G, APOBEC3C and APOBEC3H.* J Gen Virol, 2008. **89**(Pt 5): p. 1184-91.

32. Yu, Q., *et al*., *APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication.* J Biol Chem, 2004. **279**(51): p. 53379-86.

33. Kitamura, S., *et al*., *The APOBEC3C crystal structure and the interface for HIV-1 Vif binding.* Nat Struct Mol Biol, 2012. **19**(10): p. 1005-10.

34. Duggal, N.K., H.S. Malik, and M. Emerman, *The breadth of antiviral activity of Apobec3DE in chimpanzees has been driven by positive selection.* J Virol, 2011. **85**(21): p. 11361-71.

35. Bohn, M.F., *et al*., *Crystal structure of the DNA cytosine deaminase APOBEC3F: the catalytically active and HIV-1 Vif-binding domain.* Structure, 2013. **21**(6): p. 1042-50.

36. Albin, J.S., W.L. Brown, and R.S. Harris, *Catalytic activity of APOBEC3F is required for efficient restriction of Vif-deficient human immunodeficiency virus.* Virology, 2014. **450-451**: p. 49-54.

37. Siu, K.K., *et al*., *Structural determinants of HIV-1 Vif susceptibility and DNA binding in APOBEC3F.* Nat Commun, 2013. **4**: p. 2593.

38. Holden, L.G., *et al*., *Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications.* Nature, 2008. **456**(7218): p. 121-4.

39. Chen, K.M., *et al*., *Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G.* Nature, 2008. **452**(7183): p. 116-9.

40. Conticello, S.G., R.S. Harris, and M.S. Neuberger, *The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G.* Curr Biol, 2003. **13**(22): p. 2009-13.

41. Svarovskaia, E.S., *et al*., *Human apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) is incorporated into HIV-1 virions through interactions with viral and nonviral RNAs.* J Biol Chem, 2004. **279**(34): p. 35822-8.

42. Iwatani, Y., *et al*., *Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect.* J Virol, 2006. **80**(12): p. 5992-6002.

43. Byeon, I.J., *et al*., *NMR structure of human restriction factor APOBEC3A reveals substrate binding and enzyme specificity.* Nat Commun, 2013. **4**: p. 1890.

44. Marin, M., *et al*., *HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation.* Nat Med, 2003. **9**(11): p. 1398-403.

45. Mangeat, B., *et al*., *Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts.* Nature, 2003. **424**(6944): p. 99-103.

46. Zhang, J. and D.M. Webb, *Rapid evolution of primate antiviral enzyme APOBEC3G.* Hum Mol Genet, 2004. **13**(16): p. 1785-91.

47. Sawyer, S.L., M. Emerman, and H.S. Malik, *Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G.* PLoS Biol, 2004. **2**(9): p. E275.

48. OhAinle, M., *et al*., *Adaptive evolution and antiviral activity of the conserved mammalian cytidine deaminase APOBEC3H.* J Virol, 2006. **80**(8): p. 3853-62.

49. Refsland, E.W., *et al*., *Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction.* Nucleic Acids Res, 2010. **38**(13): p. 4274-84.

50. Munk, C., *et al*., *Multiple restrictions of human immunodeficiency virus type 1 in feline cells.* J Virol, 2007. **81**(13): p. 7048-60.

51. OhAinle, M., *et al*., *Antiretroelement activity of APOBEC3H was lost twice in recent human evolution.* Cell Host Microbe, 2008. **4**(3): p. 249-59.

52. Dang, Y., *et al*., *Human cytidine deaminase APOBEC3H restricts HIV-1 replication.* J Biol Chem, 2008. **283**(17): p. 11606-14.

53. Harari, A., *et al*., *Polymorphisms and splice variants influence the antiretroviral activity of human APOBEC3H.* J Virol, 2009. **83**(1): p. 295-303.

54. Tan, L., *et al*., *Sole copy of Z2-type human cytidine deaminase APOBEC3H has inhibitory activity against retrotransposons and HIV-1.* FASEB J, 2009. **23**(1): p. 279-87.

55. Modrek, B. and C. Lee, *A genomic view of alternative splicing.* Nat Genet, 2002. **30**(1): p. 13-9.

56. Zhen, A., *et al*., *A single amino acid difference in human APOBEC3H variants determines HIV-1 Vif sensitivity.* J Virol, 2010. **84**(4): p. 1902-11.

57. Li, M.M., L.I. Wu, and M. Emerman, *The range of human APOBEC3H sensitivity to lentiviral Vif proteins.* J Virol, 2010. **84**(1): p. 88-95.

58. Ooms, M., *et al*., *The localization of APOBEC3H variants in HIV-1 virions determines their antiviral activity.* J Virol, 2010. **84**(16): p. 7961-9.

59. Zhen, A., *et al*., *Reduced APOBEC3H variant anti-viral activities are associated with altered RNA binding activities.* PLoS One, 2012. **7**(7): p. e38771.

60. Wang, X., *et al*., *Analysis of human APOBEC3H haplotypes and anti-human immunodeficiency virus type 1 activity.* J Virol, 2011. **85**(7): p. 3142-52.

61. Li, M.M. and M. Emerman, *Polymorphism in human APOBEC3H affects a phenotype dominant for subcellular localization and antiviral activity.* J Virol, 2011. **85**(16): p. 8197-207.

62. Hultquist, J.F., *et al*., *Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1.* J Virol, 2011. **85**(21): p. 11220-34.

63. Bartolome, B., *et al*., *Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives.* Gene, 1991. **102**(1): p. 75-8.

64. Wyszynski, M., S. Gabbara, and A.S. Bhagwat, *Cytosine deaminations catalyzed by DNA cytosine methyltransferases are unlikely to be the major cause of mutational hot spots at sites of cytosine methylation in Escherichia coli.* Proc Natl Acad Sci U S A, 1994. **91**(4): p. 1574-8.

65. Bandaru, B., M. Wyszynski, and A.S. Bhagwat, *HpaII methyltransferase is mutagenic in Escherichia coli.* J Bacteriol, 1995. **177**(10): p. 2950-2.

66. Carpenter, M., *et al*., *Sequence-dependent enhancement of hydrolytic deamination of cytosines in DNA by the restriction enzyme PspGI.* Nucleic Acids Res, 2006. **34**(13): p. 3762-70.

67. Beletskii, A. and A.S. Bhagwat, *Correlation between transcription and C to T mutations in the non-transcribed DNA strand.* Biol Chem, 1998. **379**(4-5): p. 549-51.

68. Beale, R.C., *et al*., *Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo.* J Mol Biol, 2004. **337**(3): p. 585-96.

69. Coker, H.A., H.D. Morgan, and S.K. Petersen-Mahrt, *Genetic and in vitro assays of DNA deamination.* Methods Enzymol, 2006. **408**: p. 156-70.

70. Ooms, M., *et al*., *APOBEC3A, APOBEC3B, and APOBEC3H haplotype 2 restrict human T-lymphotropic virus type 1.* J Virol, 2012. **86**(11): p. 6097-108.

ABSTRACT

GENETIC AND BIOCHEMICAL STUDIES OF HUMAN APOBEC3H ENZYME

by

THISARI SACHITHRA ALUTHGAMA GURUGE

December 2014

Advisor: Dr. Ashok S. Bhagwat

Major: Chemistry (Biochemistry)

Degree: Master of Science

The AID/APOBEC enzymes are DNA/RNA cytosine deaminases with important functions in innate and adaptive immunity. APOBEC3 enzymes play a crucial role in restricting the replication of exogenous retroviruses such HIV-1 and endogenous retrotransposition events. In particular, APOBEC3 enzymes have evolved in humans by gene duplication to compose seven members. APOBEC3H is highly diverse in its allelic sequence and the distribution is population-specific. It has evolved under strong positive selection over millions of years. To study the catalytic mechanism, structure, function and the underlying cause of its high diversity, a high resolution crystal or NMR structure is required are not yet available. The human APOBEC3H splice variant 183 was purified as an N-terminally poly-histidine tagged construct and it was tested for biochemical activity as the first step in characterizing the enzyme.

AUTOBIOGRAPHICAL STATEMENT

Name **:** Thisari Sachithra Aluthgama Guruge

Education : MS in Chemistry (Major: Biochemistry)

Wayne State University, Detroit, Michigan, USA

BS in Biochemistry and Molecular Biology

University of Colombo, Colombo, Sri Lanka