THE ROLE OF THE RHESUS MACAQUE (*MACACA MULATTA*) APOLIPOPROTEIN B MRNA-EDITING ENZYME CATALYTIC POLYPEPTIDE-LIKE 3 (APOBEC3) IN LENTIVIRAL REPLICATION AND PERSISTENCE

BY

Kimberly Patricia Schmitt

Submitted to the graduate degree program in Anatomy and Cell Biology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Edward B. Stephens, Ph.D.
Chairperson

Nancy Berman, Ph.D.

Charlotte Vines, Ph.D.

Michael Werle, Ph.D.

Douglas Wright, Ph.D.

Date Defended: March 8, 2012
The Dissertation Committee for Kimberly P. Schmitt certifies that this is the approved version of the following thesis:

THE ROLE OF THE RHESUS MACAQUE (*MACACA MULATTA*) APOLIPOPROTEIN B MRNA-EDITING ENZYME CATALYTIC POLYPEPTIDE-LIKE 3 (APOBEC3) IN LENTIVIRAL REPLICATION AND PERSISTENCE

Edward B. Stephens, Ph.D.
Chairperson

Date Approved: March 8, 2012
II. Abstract

SHIV infections in rhesus macaques have been used to extensively study accessory proteins of HIV-1 involved in pathogenesis as well as in vaccine development. All primate lentiviruses encode for a Vif (virion infectivity factor) protein, which is required for HIV-1 replication in primary CD4$^+$ T cells and macrophages. The Vif protein interacts with APOBEC3 (A3) proteins promoting their accelerated degradation by the 26S proteosome. Sequence analysis of Vif proteins from different lentiviruses revealed that there are two highly conserved domains in the carboxyl terminus that are required for recruitment of the of the Vif-CBF-ß-Cul5/Elongin B/C/Rbx-1 E3 ubiquitin ligase complex. These domains are the viral BC box, SLQ(Y/F)LAL, and Zn$^{2+}$ binding (H-X$_5$-C-X$_{17-18}$-C-X$_{3-5}$-H; HCCH) motifs. Previous cell culture studies have shown that the introduction of amino acid substitutions into the SLQ(Y/F)LA motif resulted in decreased binding of Vif to Elongin C, while substitutions into the HCCH motif prevented the interaction of Vif with Cullin 5. In this first study, we introduced two amino acid changes in the highly conserved SLQYLA domain (S147A, L148A; AAQYLA) of the SIV Vif protein in SHIV. *In vitro*, the resulting virus, SHIV$_{VifAAQYLA}$, replicated in A3G negative CEM-SS cells but failed to replicate in A3G positive CEM cells. We also showed that hA3G was incorporated into the virion. Following these *in vitro* studies, SHIV$_{VifAAQYLA}$ was inoculated into three rhesus macaques, which were followed for over six months to assess various viral and immunological factors throughout the duration of infection. All three macaques did not develop significant CD4$^+$ T cell loss over the course of infection, had plasma viral loads that were over 100-fold lower than macaques inoculated with parental SHIV$_{KU-1bMC33}$, and developed no
histological lesions in various lymphoid tissues, and developed immunoprecipitating antibodies. DNA and RT-PCR analysis revealed that only a select number of tissues were infected with virus, while sequence analysis of PBMC and select tissue DNA (at necropsy) showed that the site-directed changes were stable during the first three weeks after inoculation but thereafter the S147A amino acid substitution changed to a threonine in two of the three macaques. However, the L148A substitution remained stable in the vif amplified from the PBMC and select tissues at necropsy in all three macaques. Extensive sequence analysis of the vif, vpu, nef, and env genes revealed a increased number of G-to-A mutations in the genes amplified from macaques inoculated with SHIVVifAAQYLA. The majority of these mutations (>85%) were in the context of 5'-TC (minus strand) and not 3'-CC, suggesting that one or more of the rhesus A3 proteins may be responsible for the observed mutational patterns. To determine if infectious virus was present in the plasma at necropsy, plasma from the three macaques inoculated with SHIVVifAAQYLA were pooled and intravenously inoculated into a naïve macaque. This macaque maintained its levels of circulating CD4+ T cells throughout the duration of infection, maintained viral loads below the limits of detection, and did not produce immunoprecipitating antibodies. However, gag was present in the DNA and RNA isolated from PBMC throughout infection and in select tissues at necropsy. The results from this first study showed for the first time the importance of the SLYQLA domain in vivo in viral pathogenesis. It also showed that mutations in vif could lead to a persistent infection in rhesus macaques resulting in the accumulation of G-to-A substitutions in the viral genome.
In the second study, we used the SHIV/macaque model of infection to compare the replication and pathogenicity of SHIVs that express a Vif protein in which the entire SLQYLA (SHIV\textsubscript{Vif5A}) or HCCH (SHIV\textsubscript{VifHCCH(-)}) domains were substituted with alanine residues. Each virus was inoculated into three rhesus macaques where various viral and immunological parameters were followed for six months. Our \textit{in vitro} results indicate that in the presence of these mutant Vif viruses, rhA3G is incorporated into the virion, stably expressed, restricts, and accumulates G-to-A substitutions (plus strand) in the \textit{nef} gene of the mutated viral genomes. \textit{In vivo}, all macaques maintained a stable level of circulating CD4\textsuperscript{+} T cells, developed low viral burdens, maintained engineered mutations, yielded no histological lesions, and developed immunoprecipitating antibodies by 12 weeks post-inoculation. However, the production of viral RNA only persisted in macaques inoculated with SHIV\textsubscript{VifHCCH(-)}. The analysis of \textit{vif} sequences amplified from PBMC DNA between weeks 0-16 during SHIV\textsubscript{VifHCCH(-)} infection revealed an increased number of G-to-A substitutions that increased with time in two of the three macaques. Sequence analysis of \textit{nef} and \textit{vpu} from the small intestine (ileum), thymus, and the spleen showed G-to-A substitutions in \textit{nef} genes isolated from macaques inoculated with SHIV\textsubscript{VifHCCH(-)}. Macaques inoculated with SHIV\textsubscript{Vif5A} effectively controlled the virus three weeks post-inoculation and no viral sequences could be amplified from tissue DNA. These studies showed that the SLQYLA and HCCH domains are critical for viral pathogenesis \textit{in vivo} and that there may exist APOBEC3 negative reservoirs in the rhesus macaque that allow for low levels of viral replication and persistence but not disease. Therefore, this study suggests that mutations targeted to one or more functional conserved domains within the Vif protein may limit viral replication and
generate an effective immune response leading to the “self-inactivation” of the virus by the activities of various APOBEC3 proteins resulting in a possible live-attenuated vaccine candidate.

The APOBEC3 family of restriction factors has been shown to inhibit certain retroviruses and retroelements. The APOBEC3 family in humans is comprised of seven cytidine deaminases (A3A, A3B, A3C, A3D, A3F, A3G, and A3H) that catalyze the deamination of cytidine to uracil on single-stranded DNA or RNA. While the human APOBEC3 repertoire has been extensively studied, the full complement of these proteins in the rhesus macaque remains unknown. Sequencing of the rhesus macaque genome has led to the identification of the rhesus homologues A3B, A3C, A3D, A3F, A3G, and A3H. Finally, we identified a human A3A (hA3A) homologue in the rhesus macaque (rhA3A) and presents evidence that both the human and rhesus Apobec3 genes are orthologous. We show that rhA3A is highly expressed in activated CD4\(^+\) T cells, widely expressed in both the visceral and central nervous system tissues of the rhesus macaque, and is degraded in the presence of the human immunodeficiency virus (HIV-1) and simian-human immunodeficiency virus (SHIV) genomes in a Vif-dependent manner. Our results also indicate that rhA3A reduced the level of infectious SHIV\(\Delta\)vif by approximately 20-fold and HIV-1\(\Delta\)vif by 3-fold. Human and monkey A3A amino acid sequences are 81% homologous and can be distinguished by a three amino acid indel located between residues 27-30. When these residues were deleted from rhA3A (rhA3A\(\Delta\)SVR), the antiviral activity of rhA3A was abolished suggesting that these residues are critical for lentivirus inhibition. Select APOBEC3 proteins are incorporated into the virion and can inhibit reverse transcription and/or
induce G-to-A hypermutation in nascent reverse transcripts in the next target cell. Previous studies revealed that rhA3G is incorporated into SHIVΔvif virions and exerts its antiviral activity in target cells by an increase in cytidine deamination of newly synthesized minus-strand viral DNA from cytosines to uracils, leading to G-to-A substitutions (plus strand) in the viral genome. We were able to detect the incorporation of rhA3A into SHIVΔvif and to a lesser extent in SHIV virions; however, we were unable to detect the incorporation of rhA3A into either HIV-1 or HIV-1Δvif virions. Even though rhA3A is incorporated into SHIVΔvif virions and potently restricts SHIVΔvif similar to rhA3G, rhA3A produced an approximately 5-fold decrease in the number of G-to-A mutations compared to rhA3G. Unlike hA3A, rhA3A did not inhibit adeno-associated virus 2 (AAV-2) replication and L1 retrotransposition. This data suggests for the first time an evolutionary switch in primate A3A virus specificity and provides evidence that a primate A3A protein can inhibit lentiviral replication.
III. Acknowledgements

First, I would like to thank my mentor Dr. Edward Stephens, for all of your help and patience. Thank you for always encouraging me to work hard, to continually challenge myself, and for introducing me to the scientific world. I would also like to thank my committee members Dr. Charlotte Vines, Dr. Nancy Berman, Dr. Michael Werle, and Dr. Doug Wright for all of your support, guidance, and time to help me finish my graduate student career. I would also like to thank Dr. Brenda Rongish for your support, guidance, and helping to keep my on track while finishing my dissertation. I would also like to extend a thank you to the Anatomy and Cell Biology department for such a great atmosphere to complete my dissertation work.

A special thank you goes out to the Stephens’ lab members: Sarah Hill, Autumn Ruiz, and Malinda Alagier. Sarah, you are an amazing person and you were an amazing help in creating experiments, bouncing around ideas, and just being there as continuous support system (“At least I figured it out”). You really helped me grow and develop as a scientist and truly are an amazing person and friend. Autumn, you are the best friend that anyone could ever want. Thank you from the bottom of my heart for supporting and encouraging me to stick through the difficult portions of graduate school. I truly wouldn’t have had the drive and ambition to complete this degree without you by my side, whether it was dancing in the lab, movies, shopping, laughing, bottles of wine, etc. Thank you for all of your scientific advice, discussions, and positive outlook on science. You really are an amazing coworker, scientist, and friend.

Mom and Mike, none of this would have been possible without your support and love. Thank you for your courage, making me laugh, making me take much needed
vacations, and inspiring me to continually challenge myself to become a better person. I would not have been able to complete this degree without you. Thank you for always believing in me even when my belief in myself has waivered. Mom, thank you for coming down to visit and take care of me during difficult times. Without your support and helping me get through surgeries and physical limitations, I would have given up. Thank you for always being there to listen to my scientific accomplishments and failures and continually encouraging me to keep pursuing my dreams. You and Mike are truly amazing and I couldn’t have asked for more supportive and loving parents.

Grandma Jean, grandma Mary, and family, thank you so much for remembering me on holidays and providing encouraging cards and gifts during my graduate studies. It always brightened my day and was truly thoughtful, thank you.

Cash and Tango, my furry little companions, thank you for always making me smile at the end of the day. Thank you for your patience while I was writing this by not sitting on my computer to distract me and always making me feel welcome when I came home. I love you both.
IV. Dedication

Dedicated to those of us who have a special place in our hearts for viruses
Table of Contents

I. **Acceptance Page** .................................................................................................................. ii

II. **Abstract** ................................................................................................................................. iii

III. **Acknowledgements** .............................................................................................................. viii

IV. **Dedication** ............................................................................................................................. x

V. **List of Figures** ......................................................................................................................... xv

VI. **Chapter One: Introduction** ..................................................................................................... 13
    A. Epidemiology ............................................................................................................................ 13
    B. Classification ............................................................................................................................. 16
    C. Geographic Distribution .......................................................................................................... 18
    D. Genomic Structure of Lentiviruses ...................................................................................... 22
    E. Viral Replication Cycle and Current HIV Therapeutic Targets ....................................... 23
    G. Viral Proteins ............................................................................................................................ 32
    H. Pathogenesis ............................................................................................................................ 56
I. Endogenous Retroelements ................................................................. 59

J. SHIV/Macaque Model of Infection ..................................................... 66

K. HIV-1 Vif Phenotype ........................................................................ 69

L. Identification of a Vif Sensitive Anti-HIV-1 Factor .......................... 70

M. The APOBEC Family of Cytidine Deaminases ............................... 71

N. Structural Characteristics of APOBEC Cytidine Deaminases ....... 74

O. Mechanism of Antiviral Activity by APOBEC3 Proteins ............... 78

P. Antiviral Activity of APOBEC3 Family of Cytidine Deaminases .... 86

Q. Vif-APOBEC3 Paradigm ................................................................. 101

R. Structural Determinants of Vif that Interact with APOBEC3 ...... 111

S. APOBEC3 and Vif as Therapeutic Targets ..................................... 113

T. Goals of the Study ........................................................................ 119

VI. Chapter Two: Mutations in the SLQYLA Motif of Vif in the Simian-Human Virus Results in a Less Pathogenic Virus in Rhesus Macaques ........ 126

A. Abstract ............................................................................................ 126

B. Introduction ...................................................................................... 127

C. Results ............................................................................................. 129

D. Discussion ....................................................................................... 167
E. Experimental Methods .................................................................171

VII. Chapter Three: Replication and Persistence of Simian-Human Immunodeficiency Viruses in Rhesus Macaques Expressing Vif Proteins with Mutations in the SLQYLA and HCCH Domains ........................................181

A. Abstract ...................................................................................181
B. Introduction .............................................................................182
C. Results ...................................................................................184
D. Discussion ...............................................................................234
E. Experimental Methods .............................................................241

VIII. Chapter Four: Differential virus restriction patterns of rhesus macaque and human APOBEC3A: Implications for lentivirus evolution .............................253

A. Abstract ...................................................................................253
B. Introduction .............................................................................253
C. Results .................................................................................................................. 256
D. Discussion ........................................................................................................... 307
E. Experimental Methods ....................................................................................... 315

IX. Chapter Five: Conclusions ............................................................................. 330

X. References .......................................................................................................... 352

XI. Appendix I .......................................................................................................... 439
    A. List of Publications .......................................................................................... 439
V. List of Figures

Chapter One: Introduction

Figure 1: Global prevalence of HIV-1 infections according to the UNAIDS 2009 report on the global AIDS epidemic.................................14-15

Figure 2: Global distribution and genetic diversity of the nine major subtypes and recombinants of HIV-1.................................................................20-21

Figure 3: Schematic representation of the genomic structures from various members of the Lentiviridae subfamily: HIV-1, SIVCPZ, HIV-2, SIVSM, and SIVMAC...24-25

Figure 4: HIV-1 replication.................................................................33-34

Figure 5: Depiction of a mature HIV-1 virion........................................35-36

Figure 6: Schematic representation of the non-LTR retrotransposon structure.....63-64

Figure 7: Schematic representation of SHIV..................................................67-68

Figure 8: APOBEC family of cytidine deaminases.........................................76-77
Figure 9: The antiviral activities of human A3G incorporated into HIV-1Δvif virions.

Table I: Summary of both the expression and restriction profile of human APOBEC3 Proteins.

Table II: Summary of the restriction profile of LTR retroelements, non-LTR retro-Elements, and other viruses by human APOBEC3 proteins.

Figure 10: Proteosomal degradation of human A3G by HIV-1 Vif.

Figure 11: Vif induced polyubiquitination of A3G through the recruitment of the CBF-β-Cul5-EloB/C-Rbx 1 E3 ubiquitin ligase complex and the specific amino acid motifs that mediate protein-protein interaction.

Figure 12: The HCCH and BC box domains of Vif are highly conserved in primate Lentiviruses.

Figure 13: The domains of HIV-1 and SIVmac239 Vif that have been implicated in binding A3 proteins.

Chapter Two: Mutations in the SLQYLA Motif of Vif in the Simian-Human Immunodeficiency Virus Results in a Less Pathogenic Virus in Rhesus Macaques
Figure 14: Replication of SHIV\textsubscript{VifAAQYLA} in permissive and nonpermissive cells…131-132

Figure 15: Incorporation of rhesus APOBEC3G into SHIV\textsubscript{VifAAQYLA} in virions……133-134

Figure 16: Circulating CD4\textsuperscript{+} T cell levels in macaques inoculated with

SHIV\textsubscript{VifAAQYLA}………………………………………………………………………………135-136

Figure 17: Plasma viral loads in macaques inoculated with SHIV\textsubscript{VifAAQYLA}………139-140

Figure 18: Stability of the SHIV\textsubscript{VifAAQYLA} Vif engineered substitution throughout

. infection………………………………………………………………………………141-142

Figure 19: Antibody response developed in macaques inoculated with

SHIV\textsubscript{VifAAQYLA}…………………………………………………………………………144-145

Figure 20. Histopathology associated with macaques inoculated with

SHIV\textsubscript{VifAAQYLA}……………………………………………………………………………146-147

Figure 21. Tissue distribution of \textit{gag} DNA in macaques inoculated with

SHIV\textsubscript{VifAAQYLA}…………………………………………………………………………149-150

Figure 22. Tissue distribution of \textit{gag} RNA in macaques inoculated with

SHIV\textsubscript{VifAAQYLA}…………………………………………………………………………151-152
Table III. Sequence analysis of viral DNA from tissue at necropsy…………………154-155

Figure 23. Representative sequence of vpu from lymphoid tissue of rhesus macaque RPL10 inoculated with SHIV_{VI\text{F}AAQLA}………………………………………156-157

Figure 24. Circulating CD4^{+} T cells and plasma viral loads of macaque I95…….159-160

Figure 25. Presence of gag in PBMC DNA at various weeks post-inoculation…161-162

Figure 26. Tissue distribution of gag DNA in macaque I95…………………………163-164

Figure 27: Tissue distribution of gag RNA in macaque I95…………………………165-166

Table IV. Oligonucleotides used in rTaq PCR to amplify SHIV genes………178-179

Chapter Three: Replication and Persistence of Simian-Human Immunodeficiency Viruses in Rhesus Macaques Expressing Vif Proteins with Mutations in the SLQYLA and HCCH Domains

Figure 28: Sequence of the SLQYLA and HCCH domain mutants………………186-187

Figure 29: Replication of SHIV_{VI\text{F}5A} and SHIV_{VI\text{F}HCCH(-)} in permissive cells and nonpermissive cells…………………………………………………188-189
Figure 30: Incorporation of rhA3G and rhA3F into SHIVVif5A and SHIV\textsubscript{VifHCCH(-)} virions…………………………………………………………………………………………………………………………190-191

Figure 31: Stability of rhA3G and rhA3F in the presence of SHIVVif5A or SHIV\textsubscript{VifHCCH(-)}…………………………………………………………………………………………………………………………193-194

Figure 32: Restriction of SHIV\textsubscript{Vif5A} and SHIV\textsubscript{VifHCCH(-)} in the presence of rhesus A3G and A3F…………………………………………………………………………………………………………………………………………196-197

Table V: \textit{In vitro} hypermutation analysis of \textit{ nef} amplified in the presence of rhesus A3G and A3F………………………………………………………………………………………………………………………………………………………………..198-199

Figure 33: Number of G-to-A substitutions obtained in \textit{ nef} in the presence of rhA3G and rhA3F…………………………………………………………………………………………………………………………………………200-202

Figure 34: Circulating CD4+ T cells levels in macaques inoculated with SHIV\textsubscript{Vif5A} and SHIV\textsubscript{VifHCCH(-)}…………………………………………………………………………………………………………………………………………..204-206

Figure 35. Plasma viral loads of macaques inoculated with SHIV\textsubscript{Vif5A} and SHIV\textsubscript{VifHCCH(-)}…………………………………………………………………………………………………………………………………………207-209

Figure 36: Sequence analysis of \textit{vif} amplified from viral DNA at various times post-
Inoculation..................................................................................................................211-213

Figure 37: Antibody responses of macaques inoculated with SHIV\textsubscript{Vif5A} and SHIV\textsubscript{VifHCCH(-)}..........................................................215-216

Figure 38: Viral copy number of \textit{gag} in the visceral tissues in macaques inoculated with SHIV\textsubscript{Vif5A} and SHIV\textsubscript{VifHCCH(-)}..........................................................218-220

Figure 39: Quantitative assessment and hypermutation plots of substitutions in the \textit{nef} gene amplified from the thymus, ileum, and spleen.................................222-225

Table VI: Percentage of G-to-A and non-G-to-A substitutions in the \textit{nef} gene amplified from the thymus, ileum, and spleen...........................................226-227

Figure 40: Quantitative assessment of substitutions in the \textit{vpu} gene amplified from the thymus, ileum, and spleen.................................................................228-231

Table VII: Percentage of G-to-A and non-G-to-A substitutions in the \textit{vpu} gene amplified from the thymus, ileum, and spleen........................................232-233

\textbf{Chapter Four:} Differential virus restriction patterns of rhesus macaque and human APOBEC3A: Implications for lentivirus evolution
Figure 41: Sequence analysis of APOBEC3A proteins.................................258-260

Figure 42: Genomic and Phylogenetic organization of rhA3A genes on
chromosome 10.........................................................................................263-265

Figure 43: Predicted sequence of rhesus APOBEC3D.................................266-267

Figure 44: Predicted sequence of rhesus APOBEC3B.................................268-269

Figure 45: Expression of rhesus A3A in visceral tissues, central nervous system
tissues, and SEB-activated CD4+ T cells.................................................272-274

Table VIII: Rhesus macaque A3 real-time PCR primers and probes..............275-276

Table VIII: Specificity evaluation of rhA3A primers and probes..................277-278

Figure 46: Expression of HA-tagged rhA3A..............................................281-283

Figure 47: Foreign DNA inhibition by rhA3A.............................................285-286

Figure 48: Rhesus A3A inhibition and incorporation.................................289-292

Figure 49: Deletion of amino acids 27-29 in rhA3A abolishes antiviral activity....294-295
Figure 50: RhA3A is sensitive to SIV and HIV-Vif mediated degradation……….296-298

Figure 51: RhA3A causes a low level of mutations in SHIVΔvif transcripts………300-302

Figure 52: RhA3A does not inhibit adeno-associated virus (AAV-2)………………..303-304

Figure 53: RhA3A does not inhibit L1 retrotransposition…………………………308-309

Chapter 5: Conclusions

Figure 54: Structural differences between human and rhesus A3A………………350-351
VI. Chapter 1: Introduction

Epidemiology

The human immunodeficiency virus type 1 (HIV-1) is the causative agent of acquired immune deficiency syndrome (AIDS) (Barre-Sinoussi et al., 1983; Gallo et al., 1984). Since its discovery in 1981, the HIV-1 pandemic has killed more than 25 million people and infects about 0.6% of the world population (UNAIDS Report on the Global AIDS Epidemic, 2010; Figure 1). By the end of 2009, the Joint United Nations Program on HIV/AIDS (UNAIDS) reports that there are approximately 33.3 million people living with HIV-1, with 22.5 million of those infected individuals residing in Sub-Saharan Africa. During 2009, there were approximately 2.6 million newly acquired infections and 1.8 million AIDS-related deaths (UNAIDS Report, 2009). The high rate of viral replication, low fidelity of reverse transcription, and frequent recombination events leads to the genetic diversity of the HIV-1 species. While there is no known cure for HIV-1 infection, the introduction of highly active anti-retroviral therapy (HAART) in 1996 has led to long lasting viral suppression reducing both the morbidity and mortality of HIV-1 infection (Palella et al., 1998). However current drugs do not eradicate HIV-1 and factors such as resistance to antiretroviral therapy occurs due to the genetic diversity of HIV-1, limited treatment options in developing countries, non-compliance, inability to afford treatments, and drug toxicity contribute to continues virus spread. Further understanding of HIV-1 cell biology and pathogenesis will be required to expand drug discovery focusing on viral/host protein interactions or the identification of cellular
Figure 1. Global prevalence of HIV-1 infections according to the UNAIDS 2009 report on the global AIDS epidemic
Adapted from UNAIDS 2009 Report on the Global AIDS Epidemic
targets conferring antiviral activity against HIV-1 thus making it possible to produce a safe and effective vaccines or antiviral treatments.

**Classification**

Simian and human immunodeficiency viruses (SIV, HIV-1, and HIV-2) are genetically related and classified in the Genus Lentivirus of the Family Retroviridae. Other members of the Genus Lentivirus include bovine immunodeficiency virus (BIV), equine infectious anemia virus (EIAV), feline immunodeficiency virus (FIV), caprine arthritis encephalitis virus (CAVE), and visna/maedi virus (MVV). Distinct SIVs has been isolated from more than 20 African nonhuman primate species, where there is extensive viremia in these monkeys without exaggerated immune response (Hahn et al., 2000; Rey-Cuille et al., 1998). Therefore, most of these particular isolates cause a nonpathogenic infection in their primary host (Hahn et al., 2000). However, several SIVs have been isolated from *Macaca nemestrina* (pig-tailed macaques; SIV*MN*) and the Asian macaque species *Macaca mulatta* (rhesus macaque; SIV*MAC*). These viruses cause an AIDS-like disease in these macaques and arose from an initially infected rhesus macaque that was being housed with an infected sooty mangabey (Hahn et al., 2000).

The extensive genetic heterogenicity of HIV-1 is often classified by the lack of proofreading by the reverse transcriptase (RT), host selective pressures, rapid turnover of the virus *in vivo*, and recombination events that occur during replication (Ho et al., 1995; Michael et al., 1999; Op de Coul et al., 2001; Roberts et al., 1988; Temin et al., 1994). HIV-1 is classified into three phylogenetic groups: group M (main), group O
(outlier), group N (non-M/O), and group P (Ayouba et al., 2000; Gurtler et al., 1994; Simon et al., 1998). Group P was recently identified from two individuals originating from Cameroon (Ayouba et al., 2001; Plantier et al., 2009; Roques et al., 2004; Simon et al., 1998; Vallari et al., 2011). From phylogenetic data we can infer that HIV-1 group M and N originated by no fewer than three independent zoonotic cross-species transmission events of chimpanzee SIV (SIV_{CPZ}; Pan troglodytes) (Gao et al., 1999; Santiago et al., 2002). Group O and P are phylogenically related to SIV_{gor} (wild-living Western lowland gorillas, Gorilla gorilla), which was derived from an ancestor of a divergent SIV_{CPZ} lineage in chimpanzees (Keele et al., 2006; Plantier et al., 2009; Takehisa et al., 2009; Van Heuverswyn et al., 2006). Of the three main groups, group M is responsible for the HIV-1 pandemic and has evolved into nine distinct phylogenetic groups: A, B, C, D, F, G, H, J, and K (Hemelaar et al., 2006). Within a given subtype there are sub-subtypes that appear to be more genetically related. These would include the subtypes A, B, and F, which are further classified into sub-subtypes A1, A2, B, B', F1, F2, respectively (Gao et al., 2001; Triques et al., 2000). The phylogenetic subtypes are currently based on either the sequence analysis of the full-length viral genome or on nucleotide sequences from multiple subgenomic regions such as gag, pol, and env. Often multiple individuals become infected with two or more subtypes to produce intersubtype recombinant forms (Peeters et al., 2001; Robertson et al., 2000). When at least three epidemiologically unlinked individuals become infected with an identical recombinant virus that has been characterized by full-length genome sequencing, the virus can be characterized as a circulating recombinant form (CRF) (Peeters et al., 2001; Robertson et al., 2000). Currently, there are 43 CRFs, with the highest diversity occurring in West Africa (Los
Alamos National Laboratory database, 2009). CRFs can be linked to areas where the parental strains are co-circulating. This increases the probability that individuals will become infected with several different genetic forms ("superinfected") resulting in the generation of unique recombinant forms (URFs) (McCutchan et al., 2006). To date, 90% of the epidemic is comprised of four subtypes (A, B, C, and D) and two CRFs (CRF01_AE and CRF02_AG) (Gilbert et al., 2007). The global distribution of these subtypes and recombinants are in Figure 2.

HIV-2, while much less wide spread than HIV-1 is classified into eight phylogenetic groups, A-H (Hemelaar et al., 2011). Unlike HIV-1, HIV-2 is thought to have originated from zoonotic transmissions from the sooty mangabey (Cercocebus atys) due to phylogenetic similarity in genomic structure (Sharp et al., 1995). It appears that seven independent transmissions from sooty mangabeys to humans occurred to result in the various HIV-2 subtypes (Chen et al., 1997; Gao et al., 1994; Yamaguchi et al., 2000).

**Geographic Distribution**

The size of the HIV-1 epidemic cause by group M, N, non-M/N varies considerably. HIV-1 group M is responsible for the majority of the global infections, approximately 33 million people. Unlike group M, group O is endemic West-Central Africa and Cameroon constituting only 1-5% of infections, group N has been found in a small group of people in Cameroon, and group P has only been identified in two individuals in Cameroon (Ayoubia et al., 2001; Plantier et al., 2009; Roques et al., 2004; Simon et al., 1998; Vallari et al., 2011). On a global scale, the most prevalent group M
subtypes are A, B, and C. Subtype C infections account for almost 50% of all worldwide HIV-1 infections and are predominantly found in >80% of infections in southern Africa and India. Subtype A viruses are predominantly located in central and eastern Africa (Kenya, Uganda, Tanzania, and Rwanda) as well as the eastern European countries formally known as the Soviet Union. The Subtype B viruses are mainly located in western and central Europe, the Americas, Australia, and in several countries of Southeast Asia, northern Africa. This subtype is also common in South African and Russian homosexual men (Buonaguro et al., 2007). The relevance of CRFs in HIV-1 infection is increasingly on the rise accounting for approximately 18% of HIV-1 infections (Hemelaar et al., 2006; Osmanov et al., 2002). CFR01-AE and CRF02-AG represent the most predominant CRFs from Southeast Asia and West/West Central Africa, respectively (Figure 2) (McCutchan et al., 1999; Montavon et al., 2000).

The HIV-2 is endemic primarily to West Africa, although there is recent spread to parts of Europe and the southwestern region of India. In Guinea-Bissau, there is prevalence up to 8-10% and Gambia is as high as 28% (Ghys et al., 1997; Langley et al., 1996; Poulsen et al., 1989; Wilkins et al., 1993). This is of concern because dual HIV-1 and HIV-2 infections are occurring more frequently in regions once considered endemic to HIV-2, which may lead to recombination events between the two viruses (Andersson et al., 1999; 2000; Ishikawa et al., 1998). The majority of HIV-2 infections are subtype A and predominant in Guinea-Bissau and Europe. To a lesser extent, HIV-2 subtype B is isolated predominantly from the eastern parts of West Africa and occasionally in Europe. More interestingly, the highest diversity of HIV-2
Figure 2. Global distribution and genetic diversity of the nine major subtypes and recombinants of HIV-1. The various colors indicate a total of ten different epidemic patterns that have been observed.
subtypes (subtypes A, B, E, and F), ~0.02% of infections, have been isolated in Sierra Leone (Reeves et al., 2002).

**Genomic Structure of Lentiviruses**

HIV-1 encodes three structural and enzymatic polyproteins; Gag (group-specific antigen), Pol (polymerase), and Env (envelope), which are analogous to proteins commonly found in all retroviruses. Additionally, the HIV-1 genome encodes two regulatory proteins; Tat (transcriptional activator), Rev (regulator of virion), as well as four accessory proteins, Nef (negative factor), Vif (virion infectivity factor), Vpr (viral protein R), and Vpu (viral protein U) (Frankel and Young, 1998). In contrast to HIV-1, HIV-2 and SIVsm phylogenetic lineages encode for an ortholog of HIV-1 Vpr known as Vpx (Viral protein X) (Goujon et al., 2007; Mailm et al., 2008; Sharova et al., 2008). Originally, HIV-2 Vpr/Vpx was viewed as the functional counterpart of HIV-1 Vpr; however, more recent evidence shows that the targeted substrate and function of these proteins differs greatly (Gibbs et al., 1994; 1995; Goujon et al., 2007; 2008; Hirsch et al., 1998; Kawamura et al., 1994; Yu et al., 1991; Mahalingam et al., 2001; Mangeot et al., 2002; Ueno et al., 2003; Wolfrum et al., 2007). Furthermore, HIV-1, SIV chimpanzees (SIVcpz), and select monkeys in the Cercopithecus genus express a Vpu homologue. However, HIV-2 and the majority of SIV lineages do not encode for a vpu gene (Barlow et al., 2003; Beer et al., 1999; 2001; McCormick-Davis et al., 2000; Strebel et al., 1998). Figure 3 illustrates the genomic structures of these members of the Genus Lentiviruses.

**Viral Replication Cycle & Current HIV Therapeutic Targets**
Using the envelope glycoprotein, gp120, HIV-1 initiates contact with the target cell by recognition of the CD4 receptor and then to a secondary chemokine coreceptor, CXCR4 or CCR5 (Figure 4) (Jakobsen et al., 2010). The CD4 receptor is a transmembrane glycoprotein expressed on the surface of T lymphocytes, T lymphocyte precursors in the bone marrow and thymus, monocytes, macrophages, dendritic cells, eosinophils, and microglial cells (Gorry and Ancuta, 2011). The high affinity binding of gp120 to CD4 triggers a conformational change in gp120 that exposes the V3 loop region of gp120 that interacts with the secondary extracellular loop of the coreceptor. A gp120-bridging sheet is formed between specific domains of gp120 after the binding of CD4 interacts with the N-terminus of the coreceptor (Moore and Nara, 1991; Chesebro et al., 1992; Wyatt et al., 1992; Milich et al., 1993; Ross et al., 1998; Hoffman and Doms, 1999; Ogert et al., 2001; Cormier and Dragic, 2002; Hoffman et al., 2002; Nabatoc et al., 2004; Rizzuto et al., 1998; Sterjovski et al., 2010). The CD4-bound gp120 with coreceptor causes further conformational changes that expose the N-terminus of gp41, which contains a fusion peptide that inserts into the host cell membrane. This occurs through a gp41 ectodomain that forms an extended coiled coil conformation and a six-helix bundle structure. This promotes the juxtaposition of the viral and host cell membrane, which is a stable structure allowing the fusion between
Figure 3. Schematic representation of the genomic structures from various members of the Lentiviridae subfamily: HIV-1, SIV<sub>CPZ</sub>, HIV-2, SIV<sub>SM</sub>, and SIV<sub>MAC</sub>. 
viral and cellular membranes resulting in a pore that allows for the release of the viral core into the host cell (Gorry and Ancuta, 2011; Helseth et al., 1991; Miller et al., 1993; Srinivas et al., 1992). There are a total of 12 different membrane proteins that have been identified as co-receptors for HIV-1. However, CXCR4 and CCR5 are the primary co-receptors for either T cell line tropic (T-tropic) or macrophage tropic (M-tropic) HIV-1 isolates, respectively (Berger et al., 1999; Deng et al., 1996; Deng et al., 1997; Doranz et al., 1996; Dragic et al., 1998; Feng et al., 1996; Gorry and Ancuta, 2011; Jakobsen et al., 2010; Liao et al., 1997). During the progression to the late stages of HIV-1 infection, 40-50% of infected individuals switch co-receptor specificity to either CXCR4 (X4 viruses) or to dual-tropic co-receptor specificity CCR5 and CXCR4 (R5/X4). The broadened cell tropism by co-receptor switching contributes to the development of a more rapid disease progression (Coetzer et al., 2008; Corry and Ancuta et al., 2011). Currently there are two anti-retrovirals that inhibit HIV-1 entry into target cells: one fusion inhibitor, enfuvirtide, and one CCR5 antagonist, maraviroc (De Clercq et al., 2009).

After the membranes fuse, the viral core is released into the cytoplasm of the cell (Figure 4). Viral uncoating is initiated with both cellular factors and the viral proteins MA, Nef, and Vif. The viral RNA genome is reverse transcribed using reverse transcriptase (RT) into a double stranded DNA copy of the viral genome (Hirsch and Curran, 1990; Harrich and Hooker, 2002). Correct regulation of the core uncoating is necessary for completion of the early stages during HIV-1 replication. TRIM5α exerts an anti-retroviral effect by targeting intact or partially uncoated viral cores (Forshey et al., 2005; Hatziiannou et al., 2004; Kar et al., 2008; Langelier et al., 2008; Sebastian
and Luban, 2005; Shi and Aiken, 2006; Stremlau et al., 2006). This block to replication occurs post-entry but pre-reverse transcription. TRIM5\(\alpha\) inhibits replication by causing rapid and premature disassembly of the viral core and by recruiting cellular proteosomal degradation machinery (Anderson et al., 2006; Campbell et al., 2008; Rold and Aiken, 2008; Perron et al., 2007; Stremlau et al., 2006; Wu et al., 2006). This factor is species specific; for example, HIV-1 is not inhibited by human TRIM5\(\alpha\) as efficiently as rhesus or owl monkey TRIM5\(\alpha\) (Bieniasz, 2004; Hatziioannou et al., 2004; Keckesova et al., 2004; Luban, 2007; Nisole et al., 2005; Ozato et al., 2008; Perron et al., 2004; Stremlau et al., 2004; Towers, 2007; Yap et al., 2004).

The reverse transcriptase (RT) is the target for three different classes of inhibitors: nucleoside RT inhibitors (NRTIs), nucleotide RT inhibitors (NtRTIs), and non-nucleoside RT inhibitors (NNRTIs). Both the NRTIs and NtRTIs interact with the substrate binding site of the RT and the NNRTIs interact with an allosteric site located a short distance away from the substrate binding site. The seven NRTIs approved for use today are zidovudine (AZT), didanosine (ddI), zalcitabine (ddC), stavudine (d4T), lamivudine (3TC), abacavir (ABC), and emtricitabine ((-)/FTC). There is one NtRTI approved for use today, tenofovir, and our NNRTIs: nevirapine, delavirdine, efavirenz, and etravirine (De Clercq, 2009).

Following uncoating the reverse transcriptase, the formation of the preintegration complex occurs (PIC) (Figure 4). This complex contains the viral reverse transcribed double-stranded DNA, matrix protein (MA), integrase (IN), viral protein R (vpr), the central DNA flap, and several host proteins (Miller et al., 1997). The PIC complex must transverse the nuclear pore complex (NPC) of the nuclear membrane to be imported into the nucleus and integrate into the host genome (Zennou et al., 2000). The NPC
spans the nuclear membrane and provides an aqueous channel allowing the bi-directional transport of proteins up to approximately 60-kDa in size (Mattaj et al., 1998; Nigg et al., 1997; Weis et al., 1998). The translocation is completed using a class of proteins known as importins and exportins. These proteins engage import and export signals within the proteins that are being mediated for transport (Nigg et al., 1997; Weis et al., 1998). The exact mechanism of the PIC-associated proteins in nuclear transportation through the NPC is unknown. However, research has shown that there is a non-classical nuclear import signal buried in the catalytic domain of IN (Bouyac-Bertoia et al., 2001). Also, there is preliminary evidence suggesting that the transit involves the association of the PIC with microtubules (Hope, unpublished data). Furthermore, a central DNA flap, which is a triple-stranded intermediate, is created during reverse transcription. The central DNA flap acts as an import signal for the PIC (Hamm et al., 1990; Zennou et al., 2000). When this structure is mutated, the HIV-1 genomes accumulated at or just inside the nuclear envelope similar to an IN nuclear localization mutant (Zennou et al., 2000). While the PIC complex is still in the cytoplasm, the viral cDNA is primed for integration. The IN trims the 3’ strand of the U3 and U5 LTRs of the viral cDNA and the ends of the viral genome in a process known as 3’-processing (3’-P). During this reaction, IN catalyzes an endonucleolytic cleavage at the 3’ site of the conserved CA dinucleotide releasing a terminal GT dinucleotide. The CA-3’-hydroxyl ends generated provide nucleophile groups used for the strand transfer. The PIC complex migrates into the nucleus as described above, and IN catalyses the insertion of the two viral cDNA ends into the host chromosome (Figure 4). During strand transfer, the two 3’-hydroxyl ends ligate to the host chromosome with a five-base pair stagger.
across the major groove DNA resulting in a two–base overhang on the viral cDNA 5’-end and a five-base pair single-stranded gap at each junction, which is repaired by host enzymes (Grandgenett, 2005; Holman and Coffin, 2005; Marcband et al., 2006; Marshall, et al., 2007; Van Maele and Debyser, 2005; Wu et al., 2005). There is only one integrase inhibitor available for clinical use, raltegravir. Raltegravir targets the strand transfer reaction. There is a similar drug now in Phase III trails that also inhibits strand transfer known as Elvitrgravir (De Chercq, 2009).

After the proviral DNA is integrated, cellular RNA polymerase II begins the first rounds of transcription of the proviral DNA. The RNA polymerase II binds cellular factors to the viral long terminal repeat (LTR) and transcription of basal amounts of Tat, Rev, and Nef begin (Jordan et al., 2001). The HIV-1 LTRs flank the viral DNA and the 3’ LTR contains three distinct regions. These regions are the unique 3’ region (U3), the repeated domain (R), and the unique 5’ region (U5). These LTR enhancer/promoter sequences are required for the replication of HIV in T cell lines. The U3 region is further divided into three regions: the core promoter, enhancer, and modulatory regions (Fields et al., 2007). The LTR also has closely related cis-acting elements that allow for promotion of transcription by RNA polymerase II. These include the TATA boxes that bind transcription factor IID, three Sp1 binding sites, enhancers containing two NF-kB binding sites, modulatory regions that contain purine box arrays, and sites that bind the Ets family members (reviewed in Kilareski et al., 2009). During the initial stage of infection, only multiply spliced mRNAs of the regulatory proteins are expressed; Tat, Nef, and Rev. Once sufficient amounts Tat are expressed (Figure 4), Tat activates further transcription of HIV-1 genes by binding the TAR element of the LTR and other
cellular transcription factors (Harrich et al, 1996; Harrich and Hooker, 2002). After sufficient amounts of Rev are expressed, Rev non-spliced and singly spliced mRNAs are produced and transported to the polysomes leading the production of other viral proteins; such as Gag, Pol, Env, Vpr, Vpu, and Vif, and genomic RNA (Pollard and Malim, 1998). Rev binds the Rev responsive element (RRE) found in singly spliced and unspliced viral RNA. This facilitates their transport through the nuclear membrane and into the cytoplasm (Sierra et al., 2005). The envelope glycoprotein precursor, gp160, is synthesized in the rough endoplasmic reticulum where high mannose moieties are added to the N-linked glycosylation sites (Asp-X-Ser/Thr). The proteins are then transported to the Golgi complex where the high mannose side chains are trimmed and complex carbohydrates are added. Additionally, the gp160 is cleaved by a host protease in the Golgi complex resulting in gp120 and the gp41 transmembrane protein. The gp120/gp41 complex proteins are transported by vesicular trafficking to the cell plasma membrane, which together with two copies of the viral RNA and the Gag and Gag-Pol precursor proteins initiate assembly (Briggs et al., 2003; Derdowski et al., 2004; Sandefur et al., 2000; Figure 4). The Gag polyprotein precursor, Pr55Gag, is essential for assembly and recruitment of components required for the formation of the virion by promoting interaction between various structural components, encapsidating the viral RNA, and associating with the envelope proteins to stimulate the budding process. Pr55Gag oligomerizes underneath the plasma membrane through protein-protein and protein-RNA interactions with the HIV-1 matrix (MA) protein facing and directly inserting into the lipid bilayer (Lee et al., 1999; Lee and Yu et al., 1998; Lingappa et al., 1997; Tritel and Resh, 2000). The budding of HIV-1 occurs by hijacking
the cellular endosomal sorting machinery (Bieniasz, 2006; Demirov and Freed, 2004; Morita and Sundquist, 2004). The p6 region of Pr55$^{Gag}$ binds to the ESCORT-1 (Endosomal Sorting Complex Required for Transport I) component Tsg101 or ALIX through the secondary assembly late domains PTAP and YPLTSL, respectively (Demirov et al., 2002; Gottlinger et al., 1991; Huang et al., 1995). ALIX binds and recruits the Snf7 subunit of ESCORT-III (McCullough et al., 2008). ESCORT-III excises the membrane neck between the viral particle and the plasma membrane. After the recruitment of all the necessary components of the virion, the immature viron buds from the plasma membrane containing the host lipid bilayer as its outer envelope (Freed, 2002; Ono and Freed, 2001). While T cells exclusively bud from the plasma membrane, HIV-1 vrions bud into vesicular structures in macrophages, which are subsequently released at the cell surface. Gag is the only protein required to form and release virion-like particles (Gelderblom, 1991; Ono, 2009; Orenstein et al., 1998). Currently, there are no anti-retrovirals that target viron assembly and budding approved for clinical trials.

The immature particle buds from the plasma membrane and the viral protease (PR) proteolytically processes by cleaving the Pr55$^{Gag}$ and the Gag-Pol precursor (Pr160$^{Gag-Pol}$) yielding cleaved Gag and Pol proteins (Adamson and Freed, 2007; Vogt, 1996). During proteolytic cleavage the virion undergoes morphological changes resulting in a spherical particle with a dense core containing the genomic RNA (Figure 5) (Ganser et al., 1999; Yeager et al., 1998). Currently, there are ten protease inhibitors approved for clinical use. These include saquinavir, ritonavir, indinavir, nelfinavir, amprenavir, lopinavir, atazanavir, fosamprenavir, tipranavir, and darunavir (De Clercq et al., 2009).
**Viral Proteins**

All of the proteins encoded by HIV-1 contribute to producing optimal conditions for HIV-1 to replicate in the host cell as well as evade the host immune system resulting in productive infection. The HIV-1 genome is flanked by two long terminal repeats (LTR) regions that do not encode for any viral proteins but are necessary for integration into the host cell genome and for encoding promoter enhancer binding sites (5’LTR).

**Gag**

The *gag* gene is expressed as a 55-kDa polyprotein precursor Pr55 that is cleaved to produce the mature HIV-1 virion. These are the matrix (MA or p17), capsid (CA or p24), and nucleocapsid (NC or p7) proteins. Additionally, the proteolytic cleavage of Pr55 results in a C-terminal domain called p6 as well as two spacer proteins known as p1 and p2 (Henderson et al., 1992; Mervis et al., 1988). The MA protein is located at the N-terminal of Pr55 and remains associated with the lipid envelope of the mature virion. Its membrane association is dependent on an N-terminal myristic acid moiety (Bryant and Ratner, 1990; Gottlinger et al., 1989; Spearman et al., 1994). A charged N-proximal region contains a cluster of conserved basic residues that bind the acidic phospholipids in the host cell lipid bilayer (Yuan et al., 1993; Zhou et al., 1994). This charged N-proximal region is essential for stabilization of Pr55 to the membrane and for extracellular particle formation and virus replication. Interaction
Figure 4. HIV-1 replication. 1) Attachment and Entry; 2) Release into the cytoplasm and uncoating; 3) Reverse Transcription; 4) Integration; 5) Transcription of proviral DNA and full-length viral genome RNA; 6) Nuclear export; 7) Processing of viral proteins; 8) Processing of Env proteins; 9) Assembly; 10) Proteolytic cleavage
Figure 5. Depiction of a mature HIV-1 virion
NIAD; Carl Henderson, 2007
between the MA and the acidic phospholipids enhance Gag membrane binding and contribute to Gag bringing the plasma membrane. A three-dimensional model structure shows that MA forms trimers with cationic charged faces resulting in Gag targeting to the anionic cytoplasmic leaflet of the plasma membrane (Hill et al., 1996; Matthews et al., 1995). An acidic phospholipid, PI(4, 5)P₂, specifically localizes to the cytoplasmic leaflet and binds to the conserved basic residue clusters on the MA proteins. This is most likely the reason Gag targets to the cytoplasmic leaflet of the plasma membrane (Chen et al., 2008; Ono et al., 2004; Saad et al., 2006). From recent studies it can be inferred that HIV-1 Gag specifically associates with raft-like microdomains in the plasma membrane (Lindwasser and Resh, 2001; Nguyen and Hildreth, 2000; Ono and Freed, 2001).

The myristylation site of Gag in the plasma membrane is generally sequestered. The insertion of this moiety into the plasma membrane is regulated by the CA protein. The CA is composed of two structurally independent, largely helical, domains: the N-terminal core and the C-terminal dimerization domain (Adamson et al., 2009; Ganser-Pornillos et al., 2008). Multimerization of the CA protein causes a conformational change in Gag allowing for the myristic acid to insert with high affinity into the plasma membrane (Saad et al., 2006). The N-terminal domain of CA also interacts with the human peptidyl-prolyl cis-trans isomerase cyclophilin A (CyPA) allowing it to be incorporated into the virion. CyPA is an intracellular receptor for the immunosuppressant cyclosporin A (CsA), which allows for optimum infectivity of HIV-1 (Franke et al., 1994; Luban et al., 1993; Handschumacher et al., 1984; Thali et al., 1994). CyPA virion incorporation stabilizes the virion and prevents human TRIM5α from
uncoating following infection (Sayah and Luban, 2004). Cleavage of CA-p2 boundary occurs late during virus maturation and is required to weaken the CA-CA interactions allowing the collapse of the spherical CA shell into a cone (Accola et al., 1998; Pettit et al., 1994; Wiegers et al., 1998).

The NC region of Pr55 contains two highly conserved copies of a CCHC-type zinc finger motif, which coordinates a zinc ion and is essential for virus replication (Aldovini and Young, 1990; Bess et al., 1992; Dorfman et al., 1993; Gorelick et al., 1993; Gorelick et al., 1990; Summers et al., 1992; Summers et al., 1990). These NC domains are also required for the packaging of the two copies of the genomic viral RNA into assembling virions (Berkowitz et al., 1996). The packaging signal (ψ) is located in the viral RNA 5' to the gag initiation codon. It is composed of four stem loop structures, which are required in cis for proper encapsidation of the viral RNA (Berkowitz et al., 1996).

Finally, the p6 domain is variable in both length and sequence. Both the PTAP (near the N-terminus) and LXXLF (Late assembly domain (L); near the C-terminus) motif are highly conserved and allow for the recruitment of the ESCORT-associated complexes. These motifs bind the ALIX (LXXLF) and TGS101 (PTAP) preventing the inward budding of vesicles into late endosomes and allow the virus to effectively bud from the plasma membrane (Bieniasz, 2006; Gottlinger et al., 1991; Huang et al., 1995; Ono and Freed, 2004; Rudner et al., 2005).

Pol
Pol is expressed from a 160-kDa Gag-pol precursor protein (Pr160Gag-Pol). This protein is synthesized following a programmed -1 ribosomal framshift that occurs at a level of 5-10% of Gag Synthesis, ensuring a strict 20:1 ratio of Gag to Gag-Pol (Jacks et al., 1998). Pr160Gag-Pol is cleaved by the protease during maturation into the viral enzymatic proteins protease (PR), reverse transcriptase (RT), and integrase (IN). The relevant aspects of PR and IN are briefly described in the overview of viral replication.

The reverse transcriptase is a multifunctional enzyme with RNA-directed DNA polymerase, DNA-directed DNA polymerase, and ribonuclease H (RNase H) activities. The crystal structure of the RT and mapping of resistance mutations has been extensively studied and is beyond the scope of this dissertation. However, it is important to understand an overview of the reverse transcription process for this dissertation. The RT reverse transcribes the single-stranded RNA from the virion into double stranded DNA, which is then integrated as provirus into the host cell. The negative sense single-stranded DNA synthesis is initiated from the tRNALys.3 bound to the primer binding site (PBS) located on the single-stranded viral RNA genome 5' end. After primer binding, the RT proceeds to the RNA 5' terminus copying repeat (R) and the unique 5' (U5) sequences. The concomitant RNase H activity hydrolyzes the ensuing RNA/DNA hybrid, which allows the complementary R sequences at the termini of the genome promote transfer of nascent DNA to the RNA 3' terminus. This allows for continued DNA synthesis. Hydrolysis of the RNA/DNA replication intermediate continues, with the exception of the 3' and central polypurine tract (cPPT). The cPPT primes the positive strand and DNA-dependent synthesis up to eighteen nucleotides of the tRNA primer. This creates the complementary positive strand PBS sequence. Following synthesis,
the PPT and tRNA primers are excised and the PBS complementary promotes a second strand transfer event. This allows for bidirectional DNA synthesis creating a double-stranded proviral DNA flanked by the LTR sequences (reviewed in Götte et al., 2010). The lack of proofreading by the RT leads to 10- to 100- fold increase in errors. The processivity of RT is also very low and recombination of DNA during DNA synthesis occurs resulting in additional mutations in the HIV-1 genome. Therefore, the errors produced by the RT, while reverse transcribing the viral genome, allow the virus to quickly evolve and become resistant to antiretroviral treatment (Jonckheer et al., 2000).

**Env**

Env is synthesized as a polyprotein, gp160, precursor from a singly spliced bicistronic vpu/env mRNA on the rough endoplasmic reticulum (RER). This precursor is cleaved at an N-terminal ER sequence that targets Env to the RER. The signal sequence is then removed by cellular signal peptidases within the ER. The gp41 transmembrane domain prevents gp160 from being fully released into the lumen of the ER. The precursor is glycosylated with N- and O-linked oligosaccharide side chains (Allan et al., 1985; Leonard et al., 1990; Bernstein et al., 1994). This allows for oligomerization of gp160 into trimers and facilitating their trafficking to the Golgi complex where it acquires high-mannose oligosaccharide side chains and complex modifications in the trans-Golgi network (Earl et al., 1990; Earl et al., 1991; Pinter et al., 1989; Schawaller et al., 1989). This is to prevent the premature interaction of oligomerized gp160 with CD4 in the secretory pathway. Since gp120 binds to the CD4 receptor on susceptible cells, the viral accessory protein Vpu binds to CD4 to downregulate its
expression by targeting it to the proteosome for ubiquitin-mediated proteasomal degradation (Fujita et al., 1997; Margottin et al., 1998; Schubert et al., 1998). In the Golgi complex, gp160 is proteolytically cleaved by cellular furin to yield the mature surface glycoprotein gp120 and the transmembrane protein gp41. This cleavage is absolutely essential for viral infectivity. Following cleavage, three molecules of each gp120 and gp41 form a heterotrimeric HIV-1 glycoprotein spike. These glycoproteins remain associated through weak, noncovalent bonds, and are transported to the surface for incorporation into the budding virion (Egan et al., 1996; Rowell et al., 1995). Env is rapidly recycled to the endosome after it arrives at the plasma membrane. This leads to low maintenance on the cell surface, and helps HIV-1 evade the host immune system (Zhu et al., 2003). The expression of gp120/gp41 complexes on the cell surface of infected cells can cause fusion of two or more cells resulting in the formation of multinucleated giant cells called syncytia (Yao et al., 1993).

The gp120 protein contains discontinuous segments five variable (V1-V5) regions, five constant (C1-C5) regions, and eighteen cysteine residues. The CD4 receptor-binding site is formed from the conserved C1, C3, and C4 domains of gp120 (Lasky et al., 1987; Kowalski et al., 1987; Olshevsky et al., 1990). The V3 loop is important for both membrane fusion and co-receptor (CXCR4 and/or CCR5) specificity (Cann et al., 1992; Chesebro et al., 1991; Freed et al., 1991; Hwang et al., 1991; O’Brien et al., 1990; Shioda et al., 1991). The dominant epitopes are found on the V3 loop and are recognized by neutralizing antibodies of the host immune system (Fouchier et al., 1992; Pollakis et al., 2004; Shioda et al., 1994). After gp120 binds to CD4, gp120 undergoes a high degree of re-folding allowing the V3 loop to bind either CXCR4 or
CCR5 through either a strong negative or strong positive charge, respectively (Doms and Peiper, 1997; Kwong et al., 2000; Moulard et al., 2000; Sullivan et al., 1998). Further conformational changes occur dissociating gp120 from gp41. Then the gp41 transmembrane glycoprotein mediates fusion of the viral lipid envelop with the host cell membrane. The gp41 subunit is composed of three major domains; an extracellular domain (or ectodomain), transmembrane domain, and a C-terminal cytoplasmic domain. The gp41 protein anchors the viral envelope to the host cell membrane. The extracellular domain contains the major fusion determinants used to catalyze the fusion of the viral and host cell lipid bilayers during viral entry (Bosch et al., 1989; Freed et al., 1990; Freed et al., 1992).

**Vif**

Vif is a basic, cytoplasmic, 23-kDa phosphoprotein expressed late in the retroviral life cycle. This protein is highly conserved among all lentiviruses (except EIAV) and is essential for viral replication. Vif is translated from a single spliced viral mRNA, whose expression is dependent on Rev (Garret et al., 1991; Schwartz et al., 1991). Vif is part of the cytoplasmic ribonucleoprotein (RNP) complex and is packaged into viral particles at low levels. Packaging of Vif occurs through interactions with viral RNA, co-packaged cellular RNA, and the nucleocapsid domain of Pr55Gag (Kahn et al., 2001; Zhang et al., 2000; Dettenhofer et al., 2000; Huvent et al., 1998; Liu et al., 1995; Karczewski et al., 1996).

Vif is required for the production of infectious virus in a cell-type specific manner. In cells termed non-permissive, viruses lacking a functional *vif* gene are restricted in
their ability to replicate and produce infectious virus, when compared to wild type virus (Courcoul et al., 1995; Fan and Peden, 1992; Fouchier et al., 1996; Gaddis et al., 2003; Goncalves et al., 1996; Hoglund et al., 1994; Simm et al., 1995; Simon and Malim, 1996; Simon et al., 1998; Sova and Volsky, 1993; von Schwedler et al., 1993). This lead to the identification of the cellular restriction factor called apolipoprotein B mRNA-editing catalytic polypeptide-like 3G (or APOBEC3G) (Sheehy et al., 2002). In the absence of Vif, APOBEC3G is encapsidated into the virion and exerts its effects on the next target cell. APOBEC3G induces massive cytidine deamination in the minus strand of viral DNA during reverse transcription producing an excessive number of G-to-A substitutions in the plus strand of DNA (Harris et al., 2003; Mangeat et al., 2003; Mariani et al., 2003; Suspene et al., 2004; Yu et al., 2004; Zhang et al., 2003). The G-to-A substitutions occur in a gradient throughout the viral genome. The most mutations are present in the Nef, and decrease towards the 5′UTR (Kijak et al., 2008; Koulinska et al., 2003; Suspene et al., 2006; 2004; Yu et al., 2004). However, Vif binds APOBEC3G and inhibits its packaging into the virion. The N-terminus of Vif binds APOBEC3G recruiting Cul5/EloB/C/Rbx-1 E2 ubiquitin ligase subsequently leading to the polyubiquitination and proteosomal degradation of APOBEC3G (Conticello et al., 2003; Kobayashi et al., 2005; Marin et al., 2003; Mehle et al., 2004; Sheehy et al., 2003; Stopak et al., 2003; Yu et al., 2003). Vif binds the E3 ubiquitin ligase complex through two highly conserved domains, $^{144}$SLQ(Y/F)LA$^{150}$ motif and the $^{108}$Hx$_5$Cx$_{17-18}$Cx$_{3-5}$H$^{139}$ domain. The $^{144}$SLQ(Y/F)LA$^{150}$ motif binds elongin C, while the $^{108}$Hx$_5$Cx$_{17-18}$Cx$_{3-5}$H$^{139}$ domain binds Cullin 5 (Luo et al., 2005; Mehle et al., 2004; 2006; Yu et al., 2004). This will be discussed in more depth below.
Nef

Nef is a small protein that is variable in length (between 200-215 amino acids) and devoid of enzymatic activity (O’Neill et al., 2006). This protein is myristoylated at the N-terminus, which is required for its expression at the plasma membrane. However, Nef is mainly localized to the paranuclear region of the infected cell with reduced expression at the plasma membrane. Nef serves as an adaptor protein to divert host cell proteins to amplify viral replication and is a major determinant of pathogenicity (Arold et al., 2001; Geyer et al., 2001; Laurent-Crawford and Hovanessian, 1992; Niederman and Ratner, 1992). In vitro, Nef has been determined to have four major functions: 1) downregulation of cell surface CD4 levels; 2) downregulation of cell surface major histocompatibility class I (MHCI) molecules; 3) mediation of cellular signaling and activation; 4) enhancement of viral particle infectivity by CD4 independent mechanisms (Aiken et al., 1996; 1997; 1998; Anderson et al., 1994; Arold et al., 2001; Aurora et al., 2000; Campbell et al., 2004; Chowers et al., 1994; Garcia et al., 1991; Greenburg et al., 1998; Lundquist et al., 2002; Luo et al., 1997; Mangasarian et al., 1997; Miller et al., 1994; Noviello et al., 2008; Pizzato et al., 2007; Renkema et al., 1999; Schwartz et al., 1996; Simmons et al., 2001; Wei et al., 2005; Wonderlich et al., 2008). Nef is defined as a pathogenic factor, because nef-deficient viruses result in lower viral burdens and severity of disease in both humans and nonhuman primates (Coleman et al., 2001; Greene et al., 2004).

The most extensively characterized function of Nef is the down-regulation of CD4 from the cell surface (Desrosiers et al., 1992; Garcia et al., 1991; Guy et al., 1987). This
involves the internalization of cell surface CD4 followed by the endosomal/lysosomal-mediated degradation. Nef binds the membrane proximal segment of the CD4 cytosolic domain and accumulates CD4 molecules into clathrin-coated pits at the cellular surface (Anderson et al., 1994; Garcia et al., 1991; Foti et al., 1997; Hua et al., 1997). A canonical dileucine motif (\textsuperscript{160}EXXXLL\textsuperscript{165}) and the two acidic residues in a structurally flexible loop that extends from amino acids 140 to 180, (\textsuperscript{174}(E/D)D\textsuperscript{175}) are required for Nef binding to adaptor protein 2 (AP-2) (Grzesiek et al., 1997; Lindwasser et al., 200). This binding facilitates the formation of clathrin-coated pits resulting in the internalization of CD4 from the cell surface and subsequent lysosomal degradation (Lindwasser et al., 2008).

Secondly, Nef allows infected cells to evade the destruction by cytotoxic T lymphocytes (CTLs) by inducing the downregulation of MHC-I molecules from the cell surface. Nef forms a ternary complex between the cytoplasmic tail of MHC-I and adaptor protein-1 (AP-1) (Noviello et al., 2008; Wonderlich et al., 2008)). AP-1 is a protein that directs proteins from the trans-Golgi network (TGN) to the endolysosomal pathways (Robinson and Bonifacino, 2001). MHC-I molecules are normally phosphorylated and transported to the plasma membrane for antigen presentation; however, the ternary complex activates a tyrosine-sorting signal in the cytoplasmic cytoplasmic tail of MHC-I. This allows the newly synthesized MHC-I molecules to be diverted from their transit to the plasma membrane to the lysosomes for degradation (Noviello et al., 2008; Roeth et al., 2006; Wonderlich et al., 2008). Nef also has been shown to increase the rate of MHC-I down-regulation from the plasma membrane by interacting with the phosphofurin acidic cluster sorting protein 1 and/or 2 (PACS-1/2).
(Atkins et al., 2008; Noviello et al., 2008; Wonderlich et al., 2008). However, much of this mechanism is still unknown. For example, we do not know how this process leads to activation of the endocytic pathway, and if Nef is blocking the recycling of MHC-I back to the plasma membrane, or if Nef delivers the MHC-I molecules to the lysosome for degradation (Foster and Garcia, 2008).

Nef regulates cellular activation through several kinases such as Pak2 and Hck (Arora et al., 2000; Briggs et al., 2000; Renkema et al., 1999; Walker et al., 1987). Although controversial, supporting data suggested that Nef regulates the actin cytoskeleton rearrangement (Fackler et al., 1999; Janardhan et al., 2004; Renkema et al., 2001; Rauch et al., 2008). Nef forms a complex with Pak2 and induces its activation early in infection. P21 activated kinases (PAK) are effectors that link Rho GTPases to cytoskeleton rearrangement and nuclear signaling (Agopian et al., 2006; Kirchhoff et al., 2004; Pulkkinn et al., 2004; Rauch et al., 2008). Nef also activates the cellular myeloid lineage specific Src family of tyrosine kinase, Hck. This protein can bind to Nef in the absence of myristylation leading Nef to form dimers and trimers (Breuer et al., 2006). However, it is unknown how Nef dimerizes and activates Hck and how this activation enhances pathogenesis.

Single cycle infection assays are used to measure how Nef enhances infectivity (Campbell et al., 2004; Pizzato et al., 2007; Qi et al., 2007; Qi et al., 2008). From these assays there are currently three models on Nef enhances infection efficiency for HIV-1 virions without the down-regulation of the CD4 receptor. First, Campbell and colleagues showed that if the actin cytoskeleton in the target cell is disrupted the infectivity of the virus was able to overcome the nef-deficient virus (Campbell et al., 2004). This shows
that the expression of Nef in the producer cell is able to overcome the cytoskeleton barrier of infection. Secondly, Pizzato and colleagues suggested that an unknown cellular protein besides CD4 blocks the function of Env. Nef downregulates the unknown protein in the producer cell, thereby blocking its incorporation into the virion and enhancing viral infectivity (Pizzato et al., 2007). Finally, several studies have shown that Nef may protect the viral core from post-fusion degradation allowing reverse transcription to continue (Qi et al., 2007; Qi et al., 2008; Wei et al., 2005).

**Vpr/Vpx**

The HIV-2 and SIV primate lentiviruses encode for a vpr gene and a paralog of vpr known as vpx gene, while HIV-1 only encodes for a vpr gene (Campbell et al., 1997). Both of these proteins are important for the transport of the pre-integration complex (PIC) into the nucleus and inducing cell cycle arrest at the G2/M phase of the cell cycle (Lang et al., 1993; Shimura et al., 1999). Both proteins have a supporting role during early infection and are selectively packaged into the virion. The incorporation occurs through a dileucine (LXXLF) motif in the p6 domain of the Gag precursor protein, Pr55Gag (Accola et al., 1999; Bachand et al., 1999; Kondo et al., 1995; Lu et al., 1995; Paxton et al., 1993; Selig et al., 1999). In addition, Vpr can regulate apoptosis and transactivate the viral promoter (Le Rouzic and Benichou, 2005; Planelles and Benichou, 2010). However, Vpx is essential for viral growth in monocyte-derived macrophages (MDMs) and primary T lymphocytes, nuclear import of viral DNA, and is critical for reverse transcription in monocyte-derived dendritic and non-dividing cells by
promoting the accumulation of full-length DNA (Fletcher et al., 1996; Fujita et al., 2008; Goujon et al., 2007; Srivastava et al., 2008).

Vpr induces cell cycle arrest at the G₂-to-M transition by activation of ataxia telangietasia-mutated and Rad-3-related kinase (ATR), which is a sensor for replication stress (Roshal et al., 2003). Vpr also associates with damaged DNA-specific binding proteins 1 (DDB1) and Cullin 4-associated factor (DCAF)-1 (Angers et al., 2006; He et al., 2006; Higa et al., 2006; Jin et al., 2006). This recruits the Cul4A\textsuperscript{DDB1} E3 ligase, which is involved in genome stability, DNA replication, and cell cycle arrest by activation of the G₂ checkpoint (DeHart and Planelles, 2008; Huang and Chen, 2008). Prolonged G₂ cell cycle arrest may attribute to Vpr-induced apoptosis; therefore, Vpr-induced cytopathicity in macrophages could be due to the absence of ATR signaling (Zimmerman et al., 2006). Also, Vpr inhibits T cell proliferation and modulates immune responses through interaction with proteins like the glucocorticoid receptor (Ayyayoo et al., 1997). One of the immune modulatory functions of Vpr is to cause a dysregulation in selective co-stimulatory molecules and cytokines in monocytes, macrophages, and dendritic cells (Granelli-Piperno et al., 2004; M et al., 2000; Majumder et al., 2005; Mirani et al., 2002; Muthumani et al., 2004; Smed-Sorensen et al., 2004).

The Vpr protein is present at the start of the replication cycle. Vpr-deficient HIV-1 shows a significant replication defect in primary macrophages, but not in T cells or laboratory cell lines (Balliet et al., 1994; Connor et al., 1995). Vpr has been shown to enhance HIV-1 transcription in monocytic cells (Varin et al., 2005). This is due to a nuclear localization signal on Vpr that interacts with proteins associated with the HIV-1 PIC, such as human uracil-DNA glycosylase (hCGI), nucleoporin, importin alpha, and
exporin 1 (Bergamaschi et al., 2009; Chen et al., 2004; Jacquot et al., 2007; Nitahara-Kasahara et al., 2007; Sherman et al., 2003). Vpr is not essential for HIV-1 infection of macrophages, but depending on the lineage of SIV, is required for replication (Bergamaschi et al., 2009; Campbell et al., 1997; Fletcher et al., 1996; Ueno et al., 2003). This shows that the vpr genes have significantly diverged and that the importance of Vpr in macrophage infection differs between lentiviral lineages.

So far, Vpx is not implicated in cell cycle arrest and does not contain a nuclear localization signal to import the PIC complex into the nucleus (Fletcher et al., 1996; Kappes et al., 1993). However, Vpx has been shown to be essential for SIVmac infection of monocytes and monocyte-derived dendritic cells (Gibs et al., 1995; Goujon et al., 2007; 2008; Wolfrum et al., 2007). In pig-tailed macaques inoculated with SIVsmPBjΔvpx, the effect of Vpx deletion was dramatic on viral pathogenicity. These macaques displayed delayed kinetics in viral replication and delayed disease manifestations (Gibbs et al., 1995; Hirsch et al., 1998). Vpx does not mimic Vpr as was previously suggested, because Vpx is required for infecting monocytes, non-diving macrophages, and monocyte-derived dendritic cells (Gibs et al., 1995; Goujon et al., 2007; 2008; Wolfrum et al., 2007). Deletion of vpr in these viruses, does not affect the requirement for Vpx in HIV-2/SIV infection (Gramberg et al., 2009; Kappes et al., 1993). However, similar to Vpr, Vpx uses the DCAF1 ubiquitin ligase to remove cellular restriction factors present in macrophages and dendritic cells (Kappes et al., 1993; Sharova et al., 2008; Srivastava et al., 2008). While the ubiquitin ligase machinery has been identified, the restriction factor degraded is still unknown.
**Tat/Rev**

Tat is a transactivator protein that contributes to trans-activation of viral and cellular genes (Ju et al., 2009; Mahlknecht et al., 2008; Nekhai et al., 2007). The HIV-1 Tat protein is synthesized during the early and late stages of viral replication and is variable in length of 86-104 amino acids. There are five functional domains in Tat, which include the N-terminus, the cysteine-rich region, the core, the basic, and the C-terminal domains. Also, there is a region known as the protein transduction domain that contains a rich-lysine and arginine region that allows Tat to cross the cell membrane (Morris et al., 2001). The extracellular form of Tat is efficient in cell membrane transduction, is released into the microenvironment by productively infected cells and, taken up by target cells in the surrounding environment (Ferrari et al., 2003; Zheng et al., 2005). Tat has several functions such as chromatin remodeling, phosphorylation of RNA polymerase II that is involved in the transcription of full-length viral mRNAs, transactivation of viral genes, and binding to nascent leader RNA known as transactivation responsive element RNA (TAR) (Ammosova et al., 2006; Richman et al., 2009; Wong et al., 2005; Yedavalli et al., 2003). Post-integration of the viral genome, the expression of the viral genome is controlled by the enhancer and promoter elements located in the (LTR) located at the 5’ end of the integrated genome (Rohr et al., 2003). To relieve the LTR, Tat interacts with several chromatin modifying complexes and histone modifying enzymes. For example, Tat recruits histone acetyltransferases (HATs) such as CREB-binding protein (CBP/p300) complex to acetylate the nucleosomes on the HIV-1 promoter region relieving LTR repression. This allows RNA polymerase II to efficiently transcribe viral genes (Deng et al., 2000; 2001; Pumfrey et al.,
By Tat changing the conformation of CBP/p300, the CBP/p300 complex is better able to bind the basal transcription factors such as the TATA-binding protein and transcription factor IIB (Dang et al., 2000; 2001). In the absence of Tat or cellular activation signals, abortive short viral mRNA transcripts are produced due to the reduced efficiency with which RNA polymerase II binds the HIV-1 promoter (Pagans et al., 2005; Williams et al., 2006). However, when Tat binds the TAR stem loop structure located at the 5' end of all initiated viral transcripts, the P-TEFb complex composed of cellular cyclin T1 and CDK9 proteins is recruited (Williams et al., 2006). This complex phosphorylates serine residues in the C-terminal domain (CTD) of RNA polymerase II leading to the synthesis of full-length HIV transcripts (Williams et al., 2006; Zhou et al., 2003). Tat also induces the phosphorylation of additional transcription factors including Sp1, CREB, the alpha subunit of eukaryotic initiation factor 2 (eIF2α), and NF-κB (Demarchi et al., 1998; Li et al., 2005; Rossi et al., 2006; Zauli et al., 2001).

Tat has been shown to modulate the expression level of other cellular proteins such as cytokines, CCR5, IL-2, and CD25, while down regulating cellular proteins such as MHC I (Matsui et al., 1996; Mayol et al., 2007; Stettner et al., 2009; Zheng et al., 2005). Since Tat can be taken up by uninfected target cells, Tat is able up regulate cellular genes to induce additional effects such as the induction of apoptosis and immunosuppression (Bennasser and Bahraoui, 2002; Gupta et al., 2008; Ju et al., 2009; Poggi and Zocci, 2006; Zheng et al., 2005). Tat can cross the blood brain barrier (BBB) and transverse cells such as neurons causing neurotoxicity, induce the release of tumor necrosis factor-α (TNF-α) from macrophages and astrocytes, IL-6 from astrocytes, and IL-1β from monocytes, causing apoptosis, which may provide an
explanation for the neurodegenerative effects of HIV-1 (Chang et al., 1997a; Chen et al., 1997; Nath et al., 1999; Cheng et al., 1998; Magnuson et al., 1995; Song et al., 2003). Tat also interacts with both non-polymerized subunits of microtubules and polymerized microtubules leading to the mitochondrial induction of apoptosis (de Mareuil et al., 2005; Giacca et al., 2005; Vendeville et al., 2004). Due to this neuronal cytoskeleton rearrangement, neuropathogenesis associated with HIV-1 infection may occur (Drewes et al., 1998).

Following integration of the provirus into the host-cell genome, the late phase of viral replication begins with transcription of the integrated proviral DNA and subsequent splicing and trafficking of viral RNAs (Grewe and Überla, 2010; Freed et al., 2001). Transcription from the LTR leads to the generation of unspliced incompletely spliced, and multiply spliced RNAs. These major types of viral transcripts are 1) full length unspliced transcript encoding the Pr55\textsuperscript{Gag} and Pr160\textsuperscript{GagPol} precursors and the full length HIV-1 ~9-kb RNA genome; 2) singly spliced RNAs that encode for Vif, Vpr, Vpu, and Nef proteins; 3) multiply spliced RNAs that encode for Tat, Rev, and Nef proteins. Rev is a 116 amino acid protein that allows mammalian cells to overcome the highly selective nuclear export of unspliced or partially spliced mRNAs (Cullen et al., 1998). The function of the Rev protein is mediated by two distinct domains: 1) an arginine-rich sequence that functions as both a nuclear localization signal (NLS) and an RNA-binding domain (RBD); and 2) a leucine-rich sequence located between amino acid residues 75 to 85 that contains the nuclear export signal (NES) (Böhnlein et al., 1991; Daly et al., 1989; Fischer et al., 1995; Hope et al., 1990; 1991; Kubota et al., 1989; Malim et al., 1989; Malim and Cullen, 1991; Malim et al., 1991; Mermer et al., 1990; Meyer et al.,
During the early stage of infection, HIV-1 full-length RNAs are expressed in the nucleus and gradually spliced to completion. These spliced RNAs are exported from the nucleus using the cellular mRNA export pathway and lead to the synthesis of Tat, Nef, and Rev. Following translation, Tat and Rev are transported back into the nucleus using nuclear localization signals (NLS) to assist in viral transcription. After enough Rev protein is expressed, nuclear export of both partially and unspliced transcripts occurs leading to the translation of all viral proteins (Cullen, 1998).

Rev function is mediated by binding a cis-acting viral RNA stem-loop structure located in the \textit{env} gene known as the Rev response element (RRE) (Malim et al., 1989). This interaction induces multimerization of Rev molecules to the RRE. Ran-GTP bound exportin 1 binds to the NES of Rev forming a complex required for the docking and translocation through the nuclear pore into the cytoplasm (Fornerod et al., 1997; Neville et al., 1997; Stade et al., 1997). Once the complex is translocated to the cytoplasm, the hydrolysis of Ran-GFP to Ran-GDP occurs inducing the release Rev from the complex. In the cytoplasm, Rev then binds importin-\(\beta\) and translocates back into the nucleus. Following translocation to the nucleus, the importin-\(\beta\)/Rev complex interacts with Ran-GTP and disassembles to release Rev back into the nucleus (Henderson et al., 1997; Palmeri et al., 1998).

\textit{Vpu}
Vpu is an 81 amino acid type I integral membrane protein. Amino acid residues 1-27 constitute the N-terminal membrane anchor followed by 54 residues that extend into the cytoplasm. The cytoplasmic component consists of a highly conserved region spanning amino acid residues 47-58 that contain a pair of serine residues that are constitutively phosphorylated by casein kinase II (Schubert et al., 1994; Tiganos et al., 1998). The cytoplasmic domain of Vpu contains two \( \alpha \)-helical structures comprised of amino acid residues 35-50 and 58-70 separated by a flexible segment containing two conserved serine residues that are phosphorylated (Federau et al., 1996; Henklein et al., 1993; Kochendoerfer et al., 2004; Wittlich et al., 2009; Wray et al., 1995; Zheng et al., 2003). HIV-1 infected cells show that Vpu interacts with several host factors such as CD4, \( \beta \)-TrCP, and BST-2. Also, homo-oligomerization of Vpu creates a pentameric complex to form cation-selective membrane pores (Ewart et al., 1996; Grice et al., 1997; Kruger and Fischer, 2009; 2010; Park et al., 2003; Schubert et al., 1996).

One function of Vpu is to induce the degradation of CD4. The Env precursor gp160 blocks the trafficking of newly synthesized CD4 by binding and forming stable complexes that are retained in the endoplasmic reticulum (ER) (Bour et al., 1991; Crise et al., 1990; Jabbar and Nayak, 1990). The cytoplasmic domain of Vpu interacts with the cytoplasmic domain of CD4 complexed with Env in the ER to target CD4 for proteosomal degradation (Bour et al., 1995; Margottin et al., 1996). The two-conserved serine residues at amino acid positions 52 and 56 in the cytoplasmic domain of Vpu are phosphorylated and interact with the WD repeats of human beta Transducin-repeat Containing Protein (\( \beta \)-TrCP) (Margottin et al., 1998; Neer et al., 1994). Human \( \beta \)-TrCP contains seven C-terminal WD repeats to mediate protein-protein interaction and an F-
box domain that acts to connect target proteins to the ubiquitin-dependent proteolytic pathway (Bai et al., 1996; Neer et al., 1994).

A second function of Vpu is the regulation of virus release from HIV-1 infected cells (Strebel et al., 1988; Terwilliger et al., 1989). There have been a number of different models that propose the regulation of virus release by Vpu. First, virus release is regulated by the Vpu-associated ion channel activity. When the transmembrane domain of Vpu is scrambled, the ion channel activity is inhibited resulting in the impairment of Vpu to regulate virus release but not CD4 degradation (Schubert et al., 1996). Secondly, Vpu may form heter-oligomeric complexes that interfere with the activity of TASK-1, a cellular ion channel (Hsu et al., 2004). In the third model, the dependence on Vpu for virus release is cell-type specific (Varthakavi et al., 2003). In addition to TASK-1, other host factors, such as BST-2, have been identified that are associated with inhibited virus release (Neil et al., 2008; Van Damme et al., 2008). BST-2 is a 180 amino acid type II transmembrane protein that was originally identified as a membrane protein in terminally differentiated human B cells of patients with multiple myeloma (Goto et al., 1994; Ohtomo et al., 1999; Ishikawa et al., 1995). This protein consists of an N-terminal transmembrane domain and a C-terminal glycosylphosphatidylinositol (GPI) anchor that allows BST-2 to associate with lipid rafts at the cell surface and on internal membranes such as the Trans-Golgi Network (TGN) (Dube et al., 2009; Kupzig et al., 2003; Masuyama et al., 2009). The ectodomain of BST-2 can form stable cysteine-linked dimmers that are critical for the inhibition of HIV-1 release (Andrew et al., 2009; Perez-Caballero et al., 2009). However, BST-2 positive cells that are infected with HIV-1ΔVpu accumulate virions at the cell surface. The BST-
2 N-terminal transmembrane domain and its C-terminal GPI anchor tethers together fully detached virions to the producer cell (Neil et al., 2008). In wild type virus, BST-2 is redistributed from the plasma membrane into early endosomes through clathrin-mediated endocytosis (Habermann et al., 2010; Neil et al., 2006; Mitchell et al., 2009; Rollason et al., 2007). Vpu interacts with BST-2 in endosomes and the TGN to downregulate BST-2 from the cell surface increasing virus release (Douglas et al., 2009; Dube et al., 2009; Habermann et al., 2010; Le Tortorex and Neil, 2009; Mitchell et al., 2009; Miyagi et al., 2009; Neil et al., 2008; Pardieu et al., 2010; Rollason et al., 2007; Sato et al., 2009; Schindler et al., 2010; Van Damme et al., 2008).

Pathogenesis

HIV-1 is transmitted through sexual contact and by direct blood-to-blood contact. The blood-to-blood contact includes intravenous drug users, occupational accidental needle sticks, transfusions, and vertical mother-to-child perinatal transmission (Curran et al., 1988; Parazzini et al., 1995). HIV-1 is the etiological agent of AIDS and results in a depletion of helper CD4+ T lymphocytes resulting in the loss of immune competence (Stevenson et al., 2003). The slow continuous depletion of CD4+ lymphocytes during the infection results in damaged immune defense leading to the susceptibility to opportunistic infections that eventually leads to death. The exact mechanism for the depletion of CD4+ lymphocytes is still unknown. There are two working hypothesis to identify the mechanism by which HIV causes disease. In the first hypothesis, it is thought that HIV causes a major depletion of CD4+ T lymphocytes by direct infection
and apoptosis. The second hypothesis considers that HIV infection is indirectly impairing cell function by a possible aberrant reaction of the immune response to the infection (Stevenson et al., 2003). A deeper understanding of viral replication, host response, reservoirs, and viral pathogenicity will help identify answers to these unanswered questions.

HIV infection can be divided into three phases: acute primary infection, asymptomatic infection, and symptomatic infection and AIDS (Coffin et al., 1995). During the acute phase of infection, ~50% of the patients remain asymptomatic, while ~50% develop flu-like symptoms such as fever, headache, sore muscles and joints, stomachache, swollen lymph glands, and/or rash, during the first four weeks of infection. During primary infection viral loads in the peripheral blood are high and the virus disseminates to many lymphoid and non-lymphoid tissues (Piatak et al., 1993). Following the acute phase, the asymptomatic carrier state begins and lasts between 10 to 12 years in an untreated typical progressor patient. During this stage, viral replication remains steady and there is a gradual decline in the number of circulating CD4+ T cells. Virus is persistently replicating in the primary and secondary lymph nodes and there is a rapid turn over of plasma virions and CD4+ T lymphocytes (Embertson et al., 1993; Pantaleo et al., 1993; Ho et al., 1995; Wei et al., 1995). This stage is also known as clinical latency. Eventually, the continuous damage to the immune system results in an increase in viral loads and life-threatening AIDS-defining diseases emerge (Sierra et al., 2005). When infected individuals establish clinical AIDS, the circulating CD4+ T lymphocytes decline to below 200 cells per microliter and/or the patient acquires an
AIDS-defining illness such as opportunistic infection(s), cancer(s), and/or HIV-related encephalopathy (Coffin et al., 1995).

Between 5-10% of patients are classified as rapid progressors and develop AIDS within 3-4 years following infection with HIV-1. These patients do not develop clinical latency, but continue to exhibit high viral loads and display a rapid decrease in the level of circulating CD4+ T lymphocytes (Arens et al., 1993; Bollinger et al., 1996; Demarest et al., 2001). There are HIV-1 patients that exhibit high viral loads initially during primary infection and then continue to present with low viral burdens for an extended period of infection. These individuals represent ~5% of the patients infected and are termed long-term non-progressors (LTNP). The level of circulating CD4+ T lymphocytes remains stable in these patients, and these individuals remain asymptomatic. These patients also display a strong qualitative and quantitative HIV-1 specific cytotoxic T-lymphocyte (CTL) response (Migueles et al., 2002). LTNP may result due to host/cellular restriction factors, attenuated HIV-1 due to mutations in nef (Δnef), or genetic polymorphisms that impair or abrogate expression of coreceptor proteins (CCR5Δ32) (Coffin et al., 1995; Deacon et al., 1995; Kirchhoff et al., 1995).

Additionally, HIV-1 invades the central nervous system early during infection and can cause persistent infection of the brain and inflammation (Shapshak et al., 2011). This infection can result in neurodegenerative disease known as neuroAIDS. The severity of this disease can range from asymptomatic neurocognitive impairment (ANI), in a mild neurocognitive disorder (MND), or be full-blown HIV-associated dementia (HAD) in infected patients (Antinori et al., 2007; Shapshak et al., 2011). These HIV-associated neurocognitive disorders (HAND) are influenced by host and viral
characteristics, co-morbid factors, substance abuse, and HARRT therapy (Bhaskaran et al., 2008). HAND is thought to occur by HIV-1 infected macrophages/monocytes transversing the tight junctions of the cerebral endothelial cells (CECs) of the blood brain barrier (BBB). Once past the BBB, the monocytes differentiate into macrophages/microglia cells (Shapshak et al., 2011). HIV-1 activates these macrophages/microglia and astrocyte cells causing the production of pro-inflammatory cytokines resulting in neuronal toxicity (Borjabad et al., 2009; Kaul et al., 2001; Minagar et al., 2004; Shapshak et al., 2004). Unfortunately, few anti-retroviral drugs treat neuroAIDS effectively because they are unable to penetrate the BBB. However, there are studies underway working to search for a way to either deliver a therapeutic drug to bypass the BBB or to create new therapeutics to treat these conditions (Shapshak et al., 2011).

**Endogenous Retroelements**

Genomic structural variations (SVs) are common in the human genome and occur through insertions, deletions, inversions, duplications, and translocations. Due to their mobility, abundance, and high sequence identity, transposable elements (TEs) are usually involved in creating genomic SVs (Konkel et al., 2010). Retrotransposons are TEs that use an RNA intermediate, are reverse transcribed, and move within the genome through a copy-and-paste mechanism (Batzer et al., 2002; Ostertag et al., 2001). Based on the presence or absence of long terminal repeats (LTRs), retrotransposons can be divided into two groups: non-LTR retrotransposons or LTR-retrotransposons (Konkel et al., 2010). Besides the non-LTR elements listed, there
are inactive old non-LTR retrotransposons encompassing approximately 6% of the human genome (Lander et al., 2001). LTR retrotransposons are the endogenous retroviruses and comprise approximately 8% of the human genome. However, there are no active LTR retrotransposones in the human genome today (Lander et al, 2001; Mills et al., 2007).

Non-LTR retrotransposons are comprised of three different families: long interspersed nuclear elements 1 (LINE1 or L1), Alu elements (a short interspersed element or SINE), and SVA (named after composite parts; SINE-R, VNTR (variable number of tandem repeats) (Cordaux et al., 2009; Belancio et al., 2008; Batzer et al., 2002; Bennett et al., 2008; Lander et al., 2001). L1s comprise about 17% (>500,000 copies) of the human genome and display evidence of ongoing activity for over the past 160 million years (Lander et al., 2001). Most L1 insertions are retrotranspositionally incompetent due to truncations and debilitating mutations; therefore, only ~80 to 100 retrotrasnposition competent L1s have been identified in the human genome (Brouha et al., 2003; Lander et al., 2001). L1s are known as autonomous retrotransposons that are currently mobilizing and at full-length are approximately 6-kb (Lander et al., 2001; Swergold et al., 1990). The element consists of a 5'UTR containing an internal RNA polymerase II (RNAPII) promoter, two open reading frames (ORF 1 and ORF 2), and a 3'UTR containing a polyadenylation signal ending with an oligo(dA)-rich tail of variable length (Figure 6) (Babushok et al., 2007; Swergold et al., 1990). ORF1 encodes for a protein involved in RNA binding and ORF2 encodes for a protein that contains both endonuclease and reverse transcriptase activity (Feng et al., 1996; Jurka et al., 1997; Martin et al., 2003; Mathias et al., 1991).
L1 retrotransposition duplication begins with RNA polymerase II transcription of the L1 gene locus from an internal promoter (Figure 6). Initiation of transcription occurs at the 5’ end of the L1 element and the internal promoter generates autonomous duplicate copies at multiple locations in the genome (Lavie et al., 2004; Swergold et al., 1990). L1 RNA is then exported to the cytoplasm where ORF1 and ORF2 are translated. ORF1 and ORF2 bind the L1 RNA transcript in a cis fashion and produce a ribonucleoprotein (RNP) complex that is transported back into the nucleus by an unknown mechanism (Lavie et al., 2004; Swergold et al., 1990; Wei et al., 2001). The integration of L1 into the genome occurs through a process referred to as target-primed reverse transcription (TPRT) (Cost et al., 2002; Feng et al., 1996; Moran et al., 1996). The L1 endonuclease cleaves the first strand of target DNA, at 5’-TTTTAA-3’ consensus sites (Jurka et al., 1997). The free 3’ hydroxyl generated by the DNA nick is used to prime the reverse transcription of L1 RNA by the L1 reverse transcriptase. The second strand of the target DNA is cleaved and used to prime second strand synthesis by unidentified mechanisms. It is thought that host repair systems are involved (Beauregard et al., 2008; Gasior et al., 2006; Gilbert et al., 2005; Ichiyanagi et al., 2007; Zingler et al., 2005). Since there are frequent staggered breaks in the host DNA at the insertion site, the retrotransposon is flanked by short stretches of host DNA between 6-20 base pairs in length and is referred to as target site duplications (TSD) (Feng et al., 1996; Szak et al., 2002). A variation of TPRT can also occur call “twin priming.” When inversions are formed, the second strand of DNA is cleaved during reverse transcription of the first strand. The 3’OH of the second strand acts as a primer for the reverse transcription on the L1 RNA and the resolution of this second cDNA produces the
inversion (Kriegs et al., 2007; Bejerano et al., 2006). Furthermore, it is thought that Alu and SVA retrotransposition occurs through the TPRT process using the L1 reverse transcriptase machinery. However, most of this process has remained elusive (Dewannieux et al., 2003; Ostertag et al., 2003; Wang et al., 2005).

There are greater than 1 million Alu copies in the human genome, making this element the most successful TE in the human genome. This element is non-autonomous and relies on the enzymatic machinery of L1 to complete retrotransposition (Wei et al., 2001; Weiner et al., 1986). Alu elements are approximately 300 base pairs long and is dimeric in structure. This element is formed by the fusion of two monomers originally derived from the 7SL RNA gene (Kriegs et al., 2007). The two monomers of Alu are divided by an A-rich linker. The 5' region contains an internal RNA polymerase III, while the 3' region contains a polyadenylation signal ending with an oligo(dA)-rich tail of a variable length (Figure 6) (Babushok et al., 2007; Swergold et al., 1990). There is no termination signal for the RNA polymerase III, so the transcripts extend into the downstream flanking sequences until a terminator can be found (Comeaux et al., 2009; Shaikh et al., 1997).

SVA elements are less characterized with approximately 3,000 copies existing in the human genome (Ostertag et al., 2003; Wang et al., 2005). This element is non-autonomous and requires the enzymatic machinery of the L1 retrotransposon. Similar to L1 elements, SVA elements are often truncated and terminate in a polyadenylation signal followed by a polyA tail (Ostertag et al., 2003; Wang et al., 2005). SVA elements are approximately 2-kb in length and are composed of a hexamer repeat region, an Alu
Figure 6. Schematic representation of the non-LTR retrotransposon structure. (A) The L1 element consists of a 5’UTR, two open reading frames (ORF1 and ORF2), and a 3’UTR. The 5’UTR contains an internal RNA polymerase II promoter (gray box). The element ends with a oligo(dA)-rich tail (AAA) preceded by a polyadenylation signal (pA). (B) The Alu element consists of two related monomers separated by an A-rich linker (consensus sequence A₅TACA₅). The left monomer consists of A and B boxes (gray boxes). These boxes are transcriptional promoters for RNA polymerase III. The 3’ end of the element contains an oligo(dA)-rich tail (AAA) that can be up to 100 base pairs long. (C) The SVA element (from 5’ to 3’) consists of a (CCTCT)ₙ hexamer repeat region, an Alu-like domain containing two antisense Alu fragments with additional sequences of unknown origin, a 35-50 base pair (VNTR) region derived from envelope polyprotein (env), and the 3’LTR of the human endogenous retrovirus (HERV)-K10. The 3’ end of this element is an oligo(dA)-rich tail preceded by a polyadenylation signal similar to that observed in (A). All three elements are flanked with target site duplications (not shown) that are generated post integration.
Adapted from Cordaux et al., 2009
like region, a VNTR (variable number of tandem repeats), a HERV-K10-like region, and a polyadenylation signal ending with a oligo(dA)-rich tail of variable length (Figure 6) (Ostertag et al., 2003; Wang et al., 2005). However, there is no internal promoter suggesting that promoter activity comes from flanking regions and that the element may be transcribed by RNA polymerase II (Ostertag et al., 2003; Wang et al., 2005).

LTR retrotransposons are very similar to exogenous retroviruses in structure. They contain two LTR’s at the 5’ and 3’ end of the element, slightly overlapping ORFs for gag, protease, and a pol gene, which encodes for the reverse transcriptase, endonuclease, integrase, and ribonuclease H (Kazazian et al., 2004; Voytas et al., 2002). LTR elements either lack or encode for a remnant of an env gene. Retrotransposition reverse transcription is essentially like retroviral reverse transcription, and the LTR retrotransposon RNA is made within the cytoplasm (Voytak et al., 2002). Some examples of LTR retrotransposons that are responsible for the majority of insertions are yeast Ty1 elements (Saccharomyces cerevisiae), IAP elements (intracisternal A-particles), MusD ETns (early transposons, and MaLR elements (mammalian LTR-retrotransposons). Most of these mobile elements are found in mice and are not present in humans, are essentially defective, and the reverse transcriptase source is still unknown (Kazazian et al., 2004k). However, IAPs and MusD elements are still highly mobile in the mouse genome (Maksakova et al., 2006). It is important to note that unlike human, endogenous retroelements in the mouse genome are 5-fold more abundant than in the human genome (Maksakova et al., 2006). In the human genome, although none of the HERVs are currently active, sequences can be recovered and reconstructed to study (Lee et al., 2007).
Disruption of coding or regulatory sequences due to deleterious insertions can occur due to non-LTR retrotransposons (Belancio et al., 2008a; Callinan et al., 2006; Chen et al., 2005; Deininger et al., 1999). This can ultimately disrupt splice sites and cause exon skipping altering gene expression and increase the risk of certain diseases like cancer (Belancio et al., 2008b; Belancio et al., 2009; Chen et al., 2008; Chen et al., 2006; Han et al., 2005; Landry et al., 2001; Perepelitsa-Belancio et al., 2003; Sobczak et al., 2002). Recent research has shown that a family of cytidine deaminases, APOBEC3 (apolipoprotein B mRNA-editing enzyme catalytic- polypeptide like 3), are capable of either introducing mutations into retroelement DNA or inhibiting retrotransposition in the absence of editing (Chen et al., 2006; Esnault et al., 2005). This is due to these retroelements undergoing a reverse transcription process analogous to exogenous retroviruses like HIV-1. These APOBEC3 proteins exude dual inhibitory effects on these endogenous retroviruses such as decreasing the number of transposed cDNA copies and extensive editing of the transposed copies (Esnault et al., 2005; Esnault et al., 2006). This material will be discussed in more detail later in the Introduction.

**SHIV/Macaque Model of Infection**

Simian-human immunodeficiency virus (SHIV) is a chimeric virus consisting of HIV-1 tat, rev, vpu, and env genes in the genetic background of pathogenic SIVmac239 (Figure 7). The pathogenic SHIV model has allowed researchers to study the role of specific genes in HIV-1 using the SHIV/macaque model of infection. Using this model, HIV-1 pathogenesis can be studied using non-human primates to evaluate mutations in
Figure 7. Schematic representation of SHIV. SIV_{mac239} genes are encoded in blue and HIV-1 genes are encoded in yellow.
HIV-1 env and vpu within the context of viral transmission, pathogenesis, and immune responses. SHIVs are known to cause AIDS and productive central nervous system (CNS) infection in rhesus macaques (Joag et al., 1997; Raghavan et al., 1997). Originally, the SHIV constructed (SHIV-4) was able to infect both rhesus and pig-tailed macaques and cause a persistent infection. However, these macaques did not develop a loss in circulating CD4+ T lymphocytes and did not develop an AIDS-like disease. SHIV-4 was passaged through a series of rhesus and pig-tailed macaques yielding pathogenic variants that developed a rapid loss of circulating CD4+ T lymphocytes, replicated to extremely high titers, and showed severe lymphoid depletion. Within eight months to a year, these macaques developed AIDS, neurological and end-stage renal disease. Also, SHIV contains unique structural motifs in gp160 of HIV-1, which allows SHIV to use the co-receptor CXCR4 similar to HIV-1 infection in humans (Joag et al., 1997; Karlsson et al., 1997; Liu et al., 1999; Shibata et al., 1997). Our laboratory uses the pathogenic SHIV variants SHIVKU-1bMC33 and SHIVKU-2MC4, which have been shown to cause a rapid loss of CD4+ T lymphocytes eventually resulting in the development of AIDS six to eight months post-infection (Liu et al., 1999; Narayan et al., 1999; McCormick-Davis et al., 2000a).

**HIV-1 Vif Phenotype**

The Vif protein is highly conserved among all lentiviruses, except for equine infectious anemia virus (EIAV), and is necessary for the production of infectious virus from CD4+ T cells and macrophages (Fisher et al., 1987; Strebel et al., 1987). Viruses lacking a functional vif gene fail to mount a spreading infection in cells termed
“nonpermissive.” These nonpermissive cells include CD4+ T cells, macrophages, and some T cell leukemia lines, such as H9, HuT78, C8166, and MT2 cells (Gabuzda et al., 1992; Sova et al., 1993; von Schwedler et al., 1993). These nonpermissive cells support the normal production of HIV-1Δvif virions, but are unable to productively infect the next target cell (Gabuzda et al., 1992; Sova et al., 1993; von Schwedler et al., 1993). Conversely, Vif is dispensable for the production of infectious virus from cells termed “permissive” (Gabuzda et al., 1992; von Schwedler et al., 1993). These permissive cells include T cell lines, such as Jurkat, SupT1, and CEM-SS cells, as well as nonhematopoietic cell lines, such as, HeLa, 293T, and COS cells (Gabuzda et al., 1992; Sova et al., 1993; von Schwedler et al., 1993). This particular “Vif phenotype” was found to be producer cell dependent, because experiments showed that transcomplementation of Vif in the producer cell was successful in rescuing viral infectivity (Strebel et al., 1987; von Schwedler et al., 1993). However, expression of Vif in the target cell had no effect (Strebel et al., 1987). Therefore, it was proposed that either Vif overcomes the effect of a negative factor in the nonpermissive cells or that permissive cells may express a host factor that influences viral infectivity.

**Identification of a Vif Sensitive Anti-HIV-1 Factor**

Using subtractive hybridization techniques, Sheehy and colleagues showed that nonpermissive cells produced an endogenous inhibitor, CEM15 or Apobec3G, which prevented the infectious spread of HIV-1 in the absence of Vif (Sheehy et al., 2002). CEM-SS (nonpermissive) cells pseudotyped with VSV-G (vesicular stomatitis virus envelope protein) were infected with HIV-1ΔvifΔenv and incubated with CEM
(permissive) cells expressing HIV-1 Env (Sheehy et al., 2002). Heterokaryons formed due to the fusion of CD4/CXCR4 receptors present on both nonpermissive and permissive cells. When progeny virions (HIV-1Δvif) were incubated with CEM-SS cells, no spreading infection resulted and HIV-1Δvif replication was inhibited (Madani et al., 1998; Simon et al, 1998; Sheehy et al., 2002). However, when the Vif protein was re-introduced into the virus, spreading infection resulted and the inhibition was relieved. Then using subtractive cloning, cDNA from the permissive (CEM) cell line was subtracted from the nonpermissive (CEM-SS) cell line to identify the cDNA clone CEM15. This Vif-sensitive antiviral cytidine deaminase is now known as APOBEC3G (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G, or hA3G). To validate its identity, cDNA from hA3G was transfected into permissive cells, and hA3G was sufficient to convert these cells into a permissive state without affecting the number of virions released (Sheehy et al., 2002).

**APOBEC Family of Cytidine Deaminases**

The APOBEC family of cytidine deaminases in human consists of activation-induced deaminase (AID), APOBEC1, APOBEC2, APOBEC3A-H, and APOBEC4 (Conticello et al., 2005; Jarmuz et al., 2002; OhAinle et al., 2006; Rogozin et al., 2005). These tissue restricted cytidine deaminases exhibit either RNA editing and/or DNA mutating activity by hydrolytic deamination of dC to dU (Harris et al., 2002; Teng et al., 1993). The gene for AID is located on human chromosome 12 and is important for humoral immune and required in activated germinal center B cells during antibody affinity maturation (Muramatsu et al., 2000). AID introduces dC to dU mutations in
single-stranded DNA substrates of specific sequence hotspots on the VDJ region of the immunoglobulin gene (Ig) by initiating somatic hypermutation (Bransteitter et al., 2003; Langlois et al., 2005; Neuberger et al., 2003; Pham et al., 2003). Hypermutation of the Ig gene enhances the affinity with which the antibody binds to and neutralizes antigens. AID also initiates class switch recombination, enabling the expression of various antibody isotypes by deamination events that lead to DNA double-stranded breaks required for the switching of constant regions of the Ig gene (Schrader et al., 2009). Individuals who lack a functional AID suffer from severe recurrent inflammatory and autoimmune disorders, whereas individuals who have abnormal AID function suffer various large B cell lymphomas and non-Hodgkin’s lymphomas (Alce et al., 2004; Luo et al., 2004; Revy et al., 2000; Svarovskaia et al., 2004).

APOBEC1 (Apo1) is also located in tandem with AID on human chromosome 12 and is primarily expressed in gastrointestinal tissue with a well-characterized role in lipid metabolism (Powell et al., 1987; Rogozin et al., 2005; Teng et al., 1993). Apo1 is the central component of an RNA editosome complex that deaminates the cytidine-6,666 in apolipoprotein B (apoB) mRNA producing two isoforms of the apoB protein (Powell et al., 1987; Rogozin et al., 2005; Teng et al., 1993). The full-length apoB100 protein mediates the transport of endogenously produced triglycerides and cholesterol in the blood (Powell et al., 1987; Yamanaka et al., 1995), and the truncated apoB48 protein regulates the absorption and transport of exogenous dietary lipids from the intestines to the tissues (Yamanaka et al., 1995). Apo1 is the most extensively studied APOBEC protein and the only APOBEC protein that requires RNA as its substrate. Over
expression of Apo1 in transgenic rabbits and mice led to both liver dysplasia and hepatocellular carcinomas (Yamanaka et al., 1995).

APOBEC2 (Apo2) is located on human chromosome 6 and is expressed specifically in skeletal muscle and heart (Jarmuz et al., 2002; Prochnow et al., 2007). Apo2 has proved to be the most enigmatic APOBEC to characterize given that its enzymatic activity is unrelated to its paralogs (Anant et al., 2001; Harris et al., 2002; Liao et al., 1999; Miki et al., 2005). Similar to Apo2, APOBEC4 (Apo4), which is located on human chromosome 1, is very poorly understood. Apo4 is primarily expressed in the testes, has low sequence similarity to the other AID/APOBEC proteins, and its function is unknown (Rogozin et al., 2005; Conticello et al., 2007). The ancestry of Apo4 reveals possible connections to tRNA-editing enzymes that may provide clues as to the origin of the APOBEC family of cytidine deaminases (Conticello et al., 2007).

APOBEC3 (A3) genes are unique to mammals and were first identified as paralogs to Apo1 (Jarmuz et al., 2002). There are a total of seven A3 genes in humans, four in felines, six in horse, two to three in artiodactyls, and a single gene in the mouse (Bogerd et al., 2008; Conticello et al., 2005; Conticello et al., 2008; LaRue et al., 2008; Munk et al., 2008). All seven A3 genes in humans are tandemly arrayed on chromosome 22 and arose through gene duplication of a single-copy primordial gene (Conticello et al., 2008; Jarmuz et al., 2002). These include APOBEC3A (A3A), APOBEC3B (A3B), APOBEC3C (A3C), APOBEC3D (A3D), APOBEC3F (A3F), APOBEC3G (A3G), and APOBEC3H (A3H) (Jarmuz et al., 2002). The A3 genes in humans demonstrate a high degree of polymorphic variation; for example, Apobec3H, which suggests that these genes are under strong selective pressure (Munk et al.,
2008; OhAinle et al., 2006; OhAinle et al., 2008; Sawyer et al., 2004; Takeda et al., 2008; Zhang et al., 2004). The A3 family of cytidine deaminases play an important role in the innate host immune response to inhibit a wide range of retroviruses, retrotransposons, hepadnaviruses, papillomavirus, foamy viruses, and parvoviruses (Abe et al., 2009; Baumert et al., 2007; Bonvin and Greeve, 2007; Bonvin et al., 2006; Chen et al., 2006; Henry et al., 2009; Jarmuz et al., 2002; Köck and Blum et al., 2008; Mahieux et al., 2005; Narvaiza et al., 2009; Noguchi et al., 2007; Paprotka et al., 2010; Sheehy et al., 2002; Strebel et al., 2005; Turelli et al., 2004; Vartanian et al., 2008; Zhang et al., 2008). However, the majority of the research on A3 proteins has been on A3G and A3F, which display the most potent antiviral activity against HIV-1 and SIV.

**Structural Characteristics of APOBEC Cytidine Deaminases**

All APOBEC cytidine deaminase family members contain a short $\alpha$-helical domain followed by a catalytic domain (CD), a short linker peptide, and a pseudocatalytic domain (PCD) (Jarmuz et al., 2002). Each catalytic domain is characterized by a highly conserved sequence required for cytidine deamination, C/H-X-E-X$_{23-28}$-P-C-X$_{2-4}$-C. The histidine and cysteine residues coordinate $\text{Zn}^{2+}$, while the critical glutamate residue is involved in the proton shuttle (Betts et al., 1994; Iwatani et al., 2006; Jarmuz et al., 2002; Wedekind et al., 2003; Xie et al., 2004). The catalytic domain of A3 proteins undergo the hydrolytic deamination at the C4 position of 2'-deoxycytidine using the conserved amino acids described above. The 2'-deoxycytidine is converted to 2'-deoxyuridine in the presence of water due to the removal of an amine group (Figure 8) (Rada et al., 2002). While there is no high-resolution crystal structure,
several structural studies of the C-terminal cytidine deaminase domain (CDA; residues 198-384) of A3G and the single domain of Apo2 have shown that the CDA core consists of five β strands flanked by α-helices on either side and appropriate connecting loops (Chen et al., 2008; Furukawa et al., 2009; Harjes et al., 2009; Holden et al., 2008; Prochnow et al., 2007; Zhang et al., 2007).

The members of the APOBEC family of cytidine deaminases contain either one or two CDA domains. Apo1, AID, Apo2, A3A, A3C, A3H, and Apo4 contain only one CDA, while A3B, A3D, A3F, and A3G contain two CDAs (Figure 8). The APOBEC proteins containing two CDAs, usually only have one domain that is catalytically active and required for species specificity, while the other CDA is involved in nucleic acid binding/sequence specificity, subcellular localization, and virion incorporation (Bogerd et al., 2007; Hache et al., 2005; Iwatani et al., 2006; Jonsson et al., 2006; Opi et al., 2006; Navarro et al., 2005; Newman et al., 2005). However, A3B is an exception because one report found both CDAs to be catalytically active (Bogerd et al., 2007).
Figure 8. APOBEC family of cytidine deaminases. A) APOBEC family members contain wither one or two CDA domain. The APOBEC proteins are aligned by their catalytically active CDA domain, depicted in purple, or their catalytically inactive CDA domain, depicted in red. The specific APOBEC protein is shown on the left and the consensus sequence is shown at the bottom. B) During deamination the 2’-deoxycytidine is converted 20 2’-deoxyuridine as a result of the removal of an amine group in the presence of water.
A. Catalytically Inactive CDA Domain
   RNA Binding & Encapsulation

<table>
<thead>
<tr>
<th>Protein</th>
<th>Catalytically Active CDA Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>APOBEC1</td>
<td></td>
</tr>
<tr>
<td>AID</td>
<td></td>
</tr>
<tr>
<td>APOBEC2</td>
<td></td>
</tr>
<tr>
<td>APOBEC3A</td>
<td></td>
</tr>
<tr>
<td>APOBEC3B</td>
<td></td>
</tr>
<tr>
<td>APOBEC3C</td>
<td></td>
</tr>
<tr>
<td>APOBEC3D</td>
<td></td>
</tr>
<tr>
<td>APOBEC3F</td>
<td></td>
</tr>
<tr>
<td>APOBEC3G</td>
<td></td>
</tr>
<tr>
<td>APOBEC3H</td>
<td></td>
</tr>
<tr>
<td>APOBEC4</td>
<td></td>
</tr>
</tbody>
</table>

H-Xaa-G-Xaa_{23-28}-P-C-Xaa_{2-4}-C

Consensus Sequence: Deaminase Active Site

B. Deamination of 2'-deoxycytidine

2'-deoxycytidine $\xrightarrow{\text{Deamination}}$ 2'-deoxyuridine
The N-terminus of double CDA A3 proteins mediate RNA binding, while the C-terminus mediates sequence specific cytidine deamination of single-stranded DNA (Friew et al., 2009; Gooch and Cullen et al., 2008; Hache et al., 2005; Iwatani et al., 2006; Li et al., 2004; Navarro et al., 2005). When mutations are introduced into the N-terminus of A3G (residues 124-127), the RNA-binding properties of A3G are impaired and A3G is no longer incorporated into budding virions (Burnett and Spearman, 2007; Huthoff et al., 2009; Huthoff and Malim, 2007; Navarro et al., 2005). The double CDA A3 proteins are also capable of both homo- and heter-oligomerization, forming higher order multimers in cells that are stabilized by RNA (Bennett et al., 2008; Burnett et al., 2007; Friew et al., 2009; Jarmuz et al., 2002; Navarro et al., 2005). For example, APOBEC3G forms homo-multimers, which do not appear to be essential for its catalytic activity or virion incorporation, but allow A3G to be form multi-protein complexes of high molecular mass (HMM) (Opi et al., 2006).

Since the deamination sequence is organized in a modular fashion at the APOBEC3 locus, the CDA domains are also characterized as either single or double Z-domain proteins (Jarmuz et al., 2002; Liddament et al., 2004). These Z-domains fall into three phylogenetic clusters depending on the conserved amino acid variations within the zinc-binding motif. These clusters are designated as type Z1, Z2, or Z3. This nomenclature is very useful in comparing cross-species and modeling the evolutionary history of the APOBEC family (Conticello et al., 2005; LaRue et al., 2008).

**Mechanism of Antiviral Activity by APOBEC3 Proteins**
In nonpermissive cells (such as activated PBMC), HIV-1ΔVif effectively incorporates approximately 7 (± 4) molecules of A3G (Xu et al., 2007). The A3G molecules are then effectively introduced into the target cell during the next round of infection and are sufficient to inhibit HIV-1 replication (Soros et al., 2007; Xu et al., 2007). In order for successful incorporation to occur, A3G directly interacts with the N-terminal region of the nucleocapsid protein (NC) of the Gag polyprotein and viral genomic RNA (Luo et al., 2004; Schafer et al., 2004; Svarovskaia et al., 2004; Zennou et al., 2004). Once inside the virion, A3G resides in the viral core, bound to the ribonucleoprotein (RNP) complex containing the viral genomic RNA, NC, integrase (IN), and Vpr proteins (Soros et al., 2007). Recently, a studied showed the importance of amino acids Y124 and W127 of A3G in facilitating RNA binding and lipid raft association. This study also showed that A3G associates with both lipid rafts and the genomic viral RNA to be effectively incorporated into the budding virion (Khan et al., 2009).

Following infection of the target cell, the minus strand of DNA is synthesized, and the RNA template is degraded by RNase H activity of the reverse transcriptase, which exposes the minus strand of DNA making it susceptible to A3G mediated degradation (Suspene et al., 2004; Yu et al., 2004). A3G then binds the single stranded DNA randomly and processively “slides or jumps” for 100 nucleotides before it begins deamination in a predominantly 3’ to 5’ manner (Chelico et al., 2006). The deamination by A3G produces excessive dC-to-dU substitutions in the minus strand of DNA (Figure 9) (Suspene et al., 2004; Yu et al., 2004). The mutated minus strand of DNA can be destroyed by two DNA base-excision repair enzymes: uracil DNA glycosylase 2 (UDG2).
or apurinic-apyrimidinic endonuclease (Schrofelbauer et al., 2005). For example, the UNG recognizes and excises the cDNA uracils, leaving an abasic site. This abasic site is cleaved by an endonuclease such as APEX1, which would cleave the cDNA backbone (Harris et al., 2003a; 2003b; Yang et al., 2007). These processes can lead to either strand breakage and impaired plus strand DNA synthesis or destruction of the minus strand of DNA (Figure 9) (Harris et al., 2003; Lecossier et al., 2003; Mangeat et al., 2003; Zhang et al., 2003). However, further studies are needed to clarify the potential mechanisms by which viral cDNA is processed and degraded following A3G mediated deamination. The few minus strands of DNA that are not degraded are synthesized into plus strand DNA where the dU promotes the misincorporation of dA resulting in G-to-A mutation (Harris et al., 2003; Mangeat et al., 2003; Mariani et al., 2003; Zhang et al., 2003). If these G-to-A substitutions exceed 10% of the dG residues in the viral genome, a phenomenon called hypermutation has occurred (Bishop et al., 2004; Harris et al., 2003; Suspene et al., 2004; Yu et al., 2004). These G-to-A substitutions occur throughout the viral genome and have been shown to produce a gradient from 3’ to 5’ in the minus strand of DNA (Kahn et al., 2007; Suspene et al., 2006). A gradient of hypermutation occurs depending on the length of time that the minus strand of remains single stranded. This is predominantly from 3’ to 5’, with the most G-to-A substitutions residing in the Env (envelope) and Nef regions and then decreasing towards the 5’ UTR (Kijack et al, 2008; Koulinska et al., 2003; Suspene et al., 2004, Suspene et al., 2006; Yu et al., 2004). Hypermutation can negate HIV-1 infectivity by leading to alteration in the viral open reading frames, the incorporation of termination codons in viral genes, impairments in plus-strand synthesis, and defects in
Figure 9. The antiviral activities of human A3G incorporated into HIV-1Δvif virions. HA3G present in the cytoplasm of cells infected with HIV-1Δvif is readily incorporated into budding virions. Following infection of the next target cell, hA3G is available to exert its antiviral effects on HIV-1 replication. These inhibitory effects on HIV-1 replication can be deaminase-dependent and/or deaminase-independent. 1) HA3G bound to viral RNA can physically impede the reverse transcriptase on the viral RNA template, thus resulting in a decrease in the number of reverse transcripts (deaminase-independent). This form of inhibition is usually incomplete and minus-stranded viral DNA is generated acting as a target for the deaminase-dependent hA3G inhibition. 2) Extensive deamination of the minus-strand of viral DNA occurs halting HIV-1 replication by either an accumulation of dG-to-dA mutations in the subsequently synthesized plus-strand of viral DNA or because the minus-strand of DNA is destroyed by uracil DNA glycosylase (UDG) and apurinic-apyrimidinic endonuclease (deaminase-dependent). 3) Also, defects in the tRNA$_{Lys}^3$ primer cleavage can lead to the formation of viral DNA with aberrant ends. This prevents the integration of the double-stranded viral DNA into the chromosome. Figure adapted from Chiu and Greene, 2008.
integration (Figure 9) (Chiu et al., 2008). This editing activity has strong antiretroviral effects and has been shown to block the replication of HIV-1, SIV, EIAV, and murine leukemia virus (MLV) (Delebecque et al., 2006; Harris et al., 2003; Mangeat et al., 2003; Mariani et al., 2003; Russell et al., 2005).

A3 proteins leave a distinct footprint of the sequence context in which they typically mutate minus sense strand viral cDNA. Since multiple A3 proteins are expressed in vivo, the sequence context helps assess which viral substrate is required for deamination. For example, A3G preferentially mutates cytidine residues in the context of 5'-CC-3' (minus strand) or 5'-GG-3' (plus strand); whereas other A3 proteins most often mutate thymidine residues resulting in 5'-TC-3' (minus strand) or 5'-GA-3' (plus strand) (Bishop et al., 2004; Harris et al., 2003; Liddament et al., 2004; Mangeat et al., 2003; Wiegand et al., 2004; Yu et al., 2004; Zhang et al., 2003; Zheng et al., 2004). Also, by assessing these sequence contexts one can determine hypermutation hotspots or determine how hypermutation can be lethal to the virus. For example, in the presence of A3G, a UGG codon encodes for a tryptophan and may be susceptible to the creation of either a UAG or even a UAA stop codon (Watts et al., 2009). Depending on the base pair next in the sequence, an AUG start codon could be replaced with a wobble base ablating the initiating methionines required for protein translation. This extensive mutation also has the ability to disrupt HIV-1 secondary RNA structures that are essential for the regulation of gene expression (Watts et al., 2009).

It is important to note that a high frequency of G-to-A mutations were evident in HIV-1 sequences from patients before the discovery of A3 proteins (Janini et al., 2001; Vartanian et al., 1991; Vartanian et al., 1994). These early studies showed that the
preferential sequence of the mutations for HIV-1 infected patients were in the context of 5'-GA-3' (plus strand) and to a lesser extent 5'-GG-3' (plus strand) (Liddament et al., 2004). However, it remains unknown which A3 protein is responsible for these reported mutations (Gandhi et al., 2008; Ulenga et al., 2008). So far, hypermutation does not appear to correlate with viral loads based off a cohort study of Senegalese female sex workers. However, this study did find that the level of hypermutation is directly correlated to level of A3G mRNA expression. This study found that patients with a lower viral set point have significantly higher levels of A3G mRNA (Ulenga et al., 2008). However, there is conflicting data on this topic. Another study involving 28 Kenyan women showed no correlation between hypermutation and viral load or CD4+ T cell counts (Piantadosi et al., 2009). Further studies are needed with larger cohorts to correlate indicators with both hypermutation and the expression of A3 proteins. Since hypermutation is common in vivo, A3 proteins through hypermutation may be aiding in the evolution of HIV-1 drug-resistance (Berkhout et al., 2004; Haché et al., 2006; Pillai et al., 2008). However, the data to support this topic is limited.

Even though the antiviral activity of A3 proteins mainly relies on the cytidine deaminase activity, A3 proteins have also been shown to wield antiviral activity independent of cytidine deamination. When mutations in the catalytic domain of A3G are introduced, deaminase activity is prevented but the mutants continue to exert antiviral activity against HIV-1 by reducing infectivity and the amount of HIV-1 cDNA produced in the target cell (Navarro et al., 2005; Newman et al., 2005). A3G has deaminase-independent functions that interfere with primer tRNA annealing, primer tRNA progression, minus- and plus- strand transfer, and DNA elongation (Anderson and
A3G contained a docking site for the C-terminus of HIV-1 integrase, which interferes with the preintegration complex (PIC) structure. This inhibits the nuclear import of the PIC complex and impairs the integration of proviral DNA (Luo et al., 2007).

A3 proteins have also been implicated in using deaminase independent functions to inhibit the replication of in retroviruses such as human T cell leukemia virus type-1 (HTLV-1), hepatitis B virus (HBV), papillomaviruses and paroviruses (Abe et al., 2009; Baumert et al., 2007; Bonvin and Greeve, 2007; Bonvin et al., 2006; Chen et al., 2006; Henry et al., 2009; Jarmuz et al., 2002; Köck and Blum, 2008; Mahieux et al., 2005; Narvaiza et al., 2009; Noguchi et al., 2007; Paprotka et al., 2010; Rosler et al., 2004; Sasada et al., 2005; Sheehy et al., 2002; Turelli et al., 2004; Vartanian et al., 2008; Zhang et al., 2008). For example, human A3G is incorporated into HTLV-1 virions and exerts an antiviral effect. No hypermutation is observed and it appears that A3G interferes with virus maturation through a deaminase-independent mechanism (Sasada et al., 2005). HBV contains a reverse transcriptase that is required for the pregenomic RNA intermediates in the cytoplasm of cells producing virus. Human A3G inhibits HBV through a deaminase-independent mechanism during the early steps of viral reverse transcription and strand elongation inhibiting viral pregenome packaging (Nguyen et al., 2007; Rosler et al., 2004; Turelli et al., 2004; Zhang et al., 2008). Human A3A inhibits the replication of parovirus adeno-associated virus (AAV) and the autonomous parovirus minute virus of mice (MVM) (Narvaiza et al., 2009). Both AAV and MVM are
single-stranded DNA viruses that do not contain a reverse transcriptase or any RNA substrates. However, hA3A inhibits the replication of both these viruses through a direct interaction with viral DNA or the replication machinery. Mutants that lack deaminase activity still retain the ability to inhibit these viruses through deaminase-independent functions (Chen et al., 2006; Narvaiza et al., 2009). Also, in the absence of hypermutation, A3A displays inhibitory activity against LINE-1 (long-interspersed nuclear element-1) and other retroelements (Bogerd et al., 2006a, 2006b; Hulme et al., 2007; Niewiadomska et al., 2007; Muckenfuss et al., 2006).

**Antiviral Activity of the APOBEC3 Family of Cytidine Deaminases**

In general, the evidence indicates that among activated Th1 CD4+ T cells, the expression of A3G and A3F are increased (Vetter et al., 2009). This increase in expression results in an increase of A3G/F incorporation into virions and an overall lower infectivity of target cells (Koning et al., 2009). This was found in most A3 proteins with the exception of A3A and to a lesser extent A3B in primary CD4+ T cells (Koning et al., 2009). IFN-α and to a lesser extent IFN-γ treatment also result in a higher expression of A3 proteins and an increased resistance to HIV-1 infection (Peng et al., 2007). IFN-α has been found to increase the expression of A3 proteins and to a greater extent A3A in macrophages. However A3B and A3C are possible exceptions to increased expression by IFN-α because they are part of the normal IFN response (Koning et al., 2009). Immature dendritic cells are resistant to HIV-1 infection due to an increased expression in A3G and A3F; however, as dendritic cells mature, the high
expression levels of these proteins are reduced (Pion et al., 2006). More work is still needed focusing on the expression of various A3 proteins by IFN.

APOBEC3G has been extensively described above; therefore, this section will mainly focus on the other six members (A3A, A3B, A3B, A3D, A3F, and A3H) of the A3 subfamily of single-stranded DNA cytidine deaminases. Compiled data regarding the inhibition of HIV-1, SIV<sub>mac</sub>239, LTR retroelements, non-LTR retroelements, hepatitis B virus, or adeno-associated virus by various human A3 proteins is listed in Table I and Table II, respectively. Of these A3 proteins, hA3B, hA3D, hA3G, and hA3F, have been shown to inhibit the replication of HIV-1Δvif, while hA3B, hA3C, hA3D, hA3G, and hA3F, have been shown to restrict SIV<sub>mac</sub>239Δvif (Dang et al., 2006; 2008; Doehle et al., 2005; Mariani et al., 2003; Wiegand et al., 2004; Yang et al., 2007; Yu et al., 2004a; 2004b; Zennou and Bieniasz, 2006; Zheng et al., 2004). Currently, more is becoming known about the rhesus A3 proteins. It is known that HIV-1Δvif can be inhibited by rhA3G, rhA3F, rhA3B, and to a lesser extent rhA3H and rhA3D, while SIV<sub>mac</sub>239Δvif has been shown to be restricted by rhA3B, rhA3D, rhA3F, rhA3G, and rhA3H (Virgen and Haziioannou, 2007; Zennou and Bieniasz, 2006).

**APOBEC3A**

Human APOBEC3A (hA3A) is expressed in the keratinocytes, spleen, small intestine, PBLs (neutrophils, primary monocytes such as macrophages and dendritic cells), bone marrow, and lungs (Refsland et al., 2010; Thielen et al., 2010). A3A is localized in both the cytoplasm and the nucleus and contains a single cytidine deaminase catalytic domain (Table I) (Goila-Gaur et al., 2007; Hultquist et al., 2011).
This A3 protein does not have antiviral activity against either HIV-1 or HIV-1Δvif (Bogerd et al., 2006; Goila-Gaur et al., 2007). When A3A is knocked down in monocytes, the susceptibility of the monocytes to HIV-1 infection increases and the spread of R5-tropic HIV-1 viruses are augmented (Berger et al., 2011; Peng et al., 2007). It is thought that this antiviral effect occurs because a large pool of hA3A exists in the target cells, that decreases the viral DNA accumulation (Berger et al., 2011). A3A is packaged into virions even in the presence of a functional Vif but does not associate with the viral nucleoprotein complex (NPC). However, when constructed into a chimeric protein with A3G, where the N-terminus of A3G was fused to A3A, the A3G-A3A chimera was packaged into the HIV-1Δvif virions by associating with the viral NPC, displayed cytoplasmic localization similar to A3G, and restored its inhibitory effects on HIV-1Δvif (Goila-Gaur et al., 2007). The target specificity sequence for hA3A has been identified as 5’-TC (minus strand) (Bulliard et al., 2011; Thielen et al., 2010). Overexpression of A3A in vitro results in increased dC-to-dU (minus strand) mutations in the human papillomavirus (HPV) genome; however it is still unknown as to whether hA3A displays deaminase-dependent restriction against hepatitis B virus (HBV) (Vartanian et al., 2008).

Another important feature of hA3A is its ability to restrict foreign DNA, which can enter cells through multiple exogenous sources. Recent research shows that the detection of foreign DNA and upregulation of IFN-α in phagocytes results in the degradation of foreign DNA by hA3A through cytidine deamination and a uracil-excision mechanism (Stenglein et al., 2010). While monocyte and in vitro cell culture genomic
DNA is unaffected, approximately 97% of cytidines in detected foreign DNA are deaminated (Stenglein et al., 2010).

A3A can block endogenous LTR-retrotransposons such as IAP and MusD as well as non-LTR retroelements such as L1 and Alu (Table II) (Bogerd et al., 2006; Chen et al., 2006). In A3A expressing cells, adeno-associated virus (AAV) producer cell replicating genomes are inhibited (Chen et al., 2006). Twelve residues (amino acids H29, K30, N57, K60, R69, F75, W98, R128, Y130, D131, E217, and P247) have been identified at the core of A3A antiviral activity and are involved in the restriction of LINE-1, AAV, and gene expression from foreign plasmid DNA. It is thought that these residues are responsible for forming the polynucleotide-accommodating groove of hA3A. If these residues are mutated the inhibition of LINE-1, AAV, and the restriction of foreign DNA is prevented by deaminase-independent mechanisms (Bulliard et al., 2011; Chen et al., 2006; Madsen et al., 1999; Melo et al., 1998; Moran et al., 1996). These mutants defective in editing lead to the accumulation of LINE-1 reverse transcripts, increasing level of viral replicative intermediates, and the expression of genes on foreign DNA (Bulliard et al., 2011; Chen et al., 2006).

Viral protein X (Vpx) has been shown to bind hA3A (Berger et al., 2010). Vpx is uniquely encoded in human immunodeficiency virus type 2 (HIV-2) and simian immunodeficiency virus (SIVmac/SIVsm) (Horton et al., 1994; Wolfrum et al., 2007). This protein is crucial for in vitro infection of human macrophages and dendritic cells, but dispensable for infection of primary lymphocytes and laboratory cell lines (Cheng et al., 2008; Fletcher et al., 1996; Sharova et al., 2008; Wolfrum et al., 2007). SIVmac Vpx has been shown to partially bind and degrade hA3A during the early stages of infection.
promoting the replication of SIV\textsubscript{mac}. If a single-point mutation in Vpx at the histidine in amino acid position 82 is substituted for an alanine, Vpx is no longer able to bind and suppress hA3A abrogating the infection of monocytes (Berger et al., 2010; Berger et al., 2011). Another study fused viral protein R (Vpr) to hA3A to convert hA3A into an inhibitor of viral replication. The Vpr protein of HIV-1 is needed to cause cell cycle arrest at the G2-to-M transition phase and to import the preintegration complexes (PIC) into the nucleus of cells (He et al., 1995; Jowett et al., 1995; Re et al., 1995; Rogel et al., 1995). The constructed fusion protein, Vpr.A3A chimera, was readily incorporated into the viral core and potently restricted both HIV-1 and SIV\textsubscript{mac} in the presence of Vif (Aguiar et al., 2008).

**APOBEC3B**

Human APOBEC3B (hA3B) is expressed in the peripheral blood leukocytes, bone marrow, lung, and stem cells (Refsland et al., 2010). A3B is not Vif sensitive and inhibits both HIV-1 and SIV\textsubscript{mac}239 in the presence of Vif, Table I (Bishop et al., 2004; Doehle et al., 2005; Yu et al., 2004). A3B contains two CDAs, as described above. Proteins that contain at double CDA display a separation of function between the domains. For example, the N-terminal CDA is responsible for viral-RNA binding and is required for encapsidation, while the C-terminal CDA determines DNA deaminase activity and substrate specificity (Bishop et al., 2004; Haché et al., 2005; Iwatani et al.,
Table I. Summary of both the expression and restriction profile by human APOBEC3 Proteins. All references for this table are listed in the text.
<table>
<thead>
<tr>
<th>Human</th>
<th>Expression Profile</th>
<th>Subcellular Localization</th>
<th>HIV-1</th>
<th>HIV-1 ΔVif</th>
<th>SIVmac</th>
<th>SIVmac ΔVif</th>
</tr>
</thead>
<tbody>
<tr>
<td>APOBEC3A</td>
<td>Keratinocytes, spleen, small intestine, PBLs, bone marrow, lung</td>
<td>Cytoplasm &amp; Nucleus</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>APOBEC3B</td>
<td>Keratinocytes, colon, small intestine, testis, ovaries, stem cells, PBLs, bone marrow, lung</td>
<td>Nuclear</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>APOBEC3C</td>
<td>PBLs, thymus, spleen, lymph node, testis, ovary, small intestine, colon, liver, pancreas, heart, lung, adipose tissue</td>
<td>Cytoplasm &amp; Nucleus</td>
<td>No</td>
<td>Weak</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>APOBEC3D</td>
<td>PBLs, thymus, spleen, lymph node, ovary, liver, lung, adipose</td>
<td>Cytoplasm &amp; Nucleus</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>APOBEC3F</td>
<td>PBLs, thymus, spleen, lymph node, testis, ovary, uterus, brain, lung, colon, liver, kidney, pancreas</td>
<td>Cytoplasm</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>APOBEC3G</td>
<td>PBLs, thymus, spleen, lymph node, testis, ovary, uterus, brain, lung, small intestine, colon, liver, kidney, pancreas</td>
<td>Cytoplasm</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>APOBEC3H</td>
<td>PBLs, thymus, testis, ovary, brain, small intestine, colon, keratinocytes</td>
<td>Cytoplasm &amp; Nucleus</td>
<td>No</td>
<td>Select Variants</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
Table II. Summary of the restriction profile of LTR retroelements, non-LTR retroelements, and other viruses by human APOBEC3 Proteins. All references used in this table can be found in the text.
<table>
<thead>
<tr>
<th>Human</th>
<th>IAP</th>
<th>MusD</th>
<th>Ty1</th>
<th>L1</th>
<th>Alu</th>
<th>HBV</th>
<th>AAV</th>
<th>HPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>APOBEC3A</td>
<td>Yes</td>
<td>Yes</td>
<td>?</td>
<td>Yes</td>
<td>Yes</td>
<td>?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>APOBEC3B</td>
<td>Yes</td>
<td>Yes</td>
<td>?</td>
<td>Yes</td>
<td>Yes</td>
<td>Moderate</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>APOBEC3C</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>APOBEC3F</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>?</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>APOBEC3G</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>APOBEC3H</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>Yes, only stably expressed HaplI</td>
<td>Yes, only stably expressed HaplII</td>
<td>Yes, only stably expressed HaplII</td>
<td>?</td>
<td>Yes, only stably expressed HaplII</td>
</tr>
</tbody>
</table>
2006; Liddament et al., 2004; Navarro et al., 2005; Newman et al., 2005; Wiegand et al., 2004). However, the C-terminal CDA of A3B is catalytically dominant and determines sequence specificity (5'-TC, minus strand), but the N-terminal CDA also elicits deaminase activity that is capable of mutating the viral genome (Bogerd et al., 2008). Human A3B has also been found to inhibit both LTR and non-LTR retroelements (Table II) (Bo derg et al., 2006; Chen et al., 2006; Esnault et al., 2005; Muskenfuss et al., 2006; Stenglein et al., 2006). Mouse LTR retrotransposons, IAP and MusD elements are strongly inhibited by hA3B (Bogerd et al., 2006; Chen et al., 2006; Esnault et al., 2005).

Secondly, hA3B strongly inhibits non-LTR retorelements such as L1 and Alu (Bogerd et al., 2006; Chen et al., 2006; Muckenfuss et al., 2006; Stenglein et al., 2006). Using CDA mutants that lack deaminase activity, A3B still effectively inhibits L1 retrotransposition by approximently 75% (Wissing et al., 2011). This may occur due to a highly homologous region in the CDA that directly interacts with L1 ORF2 (Turelli et al., 2004). A3B is predominantly localized to the nucleus, and in overexpression experiments performed in HeLa cells the nuclear localization of A3B is not essential for A3B to inhibit L1 retrotransposition (Hultquist et al., 2011; Pak et al., 2011). Finally, in cotransfected hepatoma cell lines, hA3B used deaminase-independent functions to interfere with the replication of HBV (Okeoma et al., 2007). This is due to hA3B having two splice variants in the human liver, which are both expressed, but only the longer variant is capable to inhibit HBV (Bonvin et al., 2006).
Interestingly, a small ~4-kb deletion in A3B is found in some patients with breast cancer (Komatsu et al., 2008). Also, some patients display homozygous deletions of A3B, which can be associated with malignant tissues (Iskow et al., 2010). There is also a common deletion polymorphism in the Oceanic population where 29.5 kb of genomic sequence spanning from the fifth exon of A3A to the eighth exon of A3B. While A3B is deleted, A3A is produced and is a functional-full length A3A fusion protein with the 3'UTR of A3B (Kidd et al., 2007). At this time, it is unknown if this fusion protein is functional, restricts L1 retrotransposition, or what the potential role of this protein could be.

**APOBEC3C**

Human APOBEC3C (hA3C) is insensitive to IFN and expressed in the PBLs, thymus, spleen, lymph node, testis, ovary, small intestine, colon, liver, pancreas, heart, lung, and adipose tissue (Table I) (Refsland et al., 2010). A3C contains one CDA and is localized to both the cytoplasm and the nucleus (Hultquist et al., 2011). In contrast to A3B, the deaminase activity of A3C is poorly active against its sequence specific target, 5'-TC (minus strand), and requires dimerization to exert its antiviral effect (Stauch et al., 2009; Yu et al., 2004). This A3 protein has potent antiviral activity against SIV\textsubscript{mac}239 and to a much lesser extent HIV-1\Delta\textit{vif} (Yu et al., 2004) (Table I). Even though A3C is efficiently incorporated into the virions of these viruses, very few G-to-A (plus strand) mutations result (Yu et al., 2004). If the N-terminal region (amino acid residues 2-47, before the CDA) of hA3C is deleted, than hA3C is insensitive to HIV-1 or SIV\textsubscript{agm} (African green monkey) Vif (Zhang et al., 2008). This indicates that this region contains
an important determinant required for Vif-induced degradation. When residue E106, located after the CDA, is mutated, hA3C is no longer sensitive to HIV-1 Vif (Smith et al., 2010). Also, the amino acid residue E289 in hA3C has been shown to be critical for HIV-1 Vif mediated degradation (Smith et al., 2010).

When hA3C is overexpressed in vitro cell culture, hA3C mildly inhibits Non-LTR Retroelements such as L1 and Alu (Bogerd et al., 2006; Muckenfuss et al., 2006; Stenglein et al., 2006). In contrast, hA3C effectively inhibits LTR retroelements such as IAP and MusD (Bogerd et al., 2006; Chen et al., 2006; Esnault et al., 2005) (Table II). In tissue culture, hA3C overexpression can reduce herpes simplex virus 1 (HSV-1) and Epstein-Barr herpesvirus (EBV) viral titers, decrease the particle/plaque forming unit ratio by approximately 10-fold, and hypermutate genomes (Suspène et al., 2011a). Human A3C has also been shown to mutate transfected human papillomavirus (HPV) and edit mitochondrial DNA (mtDNA) (Suspène et al., 2011b; Vartanian et al., 2008). Excessive hypermutation of the hepatitis B virus (HBV) genome has also been reported (Baumert et al., 2007; Köck et al., 2008).

**APOBEC3D**

Human APOBEC3D (hA3D; formally known as APOBEC3DE) is expressed in the PBLs, thymus, spleen, lymph node, ovary, liver, lung, adipose tissue, and more extensively than hA3F in nonpermissive cells (Dang et al., 2006; Refsland et al., 2010). This A3 protein is localized similar to A3C in both the cytoplasm and nucleus but contains two CDAs (Hultquist et al., 2011). HA3D inhibits the replication of both HIV-1Δvif and SIVmac239Δvif (Table I) (Dang et al., 2006). This A3 protein is efficiently
incorporated into virions in the absence of Vif, which resulted in excessive G-to-A hypermutation. A3D introduces 5’-CC, 5’-TC, and 5’-GC (minus strand shown) mutations into newly synthesized viral minus-strand DNA (Dang et al., 2006). Using a GST-hA3D pull down assay in cells expressed with hA3G and hA3F, hA3D formed heteromultimers with both A3G and A3F which may enhance the antiretroviral effect (Dang et al., 2006; Wiegand et al., 2004). Very little is known about the inhibition of hA3D on other viruses, LTR retrotransposons, or non-LTR retrotransposons. Therefore, more studies are needed to determine the extent to which hA3D exerts its antiviral activities.

**APOBEC3F**

Human APOBEC3F (hA3F) is expressed in the PBLs, thymus, spleen, lymph node, testis, ovary, uterus, brain, lung, colon, kidney, and pancreas (Refsland et al., 2010). The cellular localization of hA3F is cytoplasmic (Table I) (Hultquist et al., 2011). HA3F is the second most potent inhibitor of HIV-1Δvif, but is produced at lower levels in natural target cells of HIV-1, is partially resistant to Vif, and is less potent than hA3G (Koning et al., 2009; Liddament et al., 2004; Simon et al., 2005; Xu et al., 2004; Wiegand et al., 2004; Zennou and Bieniasz, 2006). Mutations in binding domain of Vif that prevent interaction with hA3G increase the sensitivity of Vif to hA3F (Russell et al., 2009). This protein has two catalytically active cytidine deaminase domains; however, unlike A3G, the deaminase activity in the C-terminal CDA is not required for inhibition of HIV-1Δvif. Catalytically inactive mutants of A3F are just as potent inhibitors of viral infectivity causing reductions in the accumulation of viral reverse transcripts (Holmes et
Therefore, A3F-mediated inhibition of HIV-1Δvif infectivity primarily occurs through a deaminase-independent mechanism (Holmes et al., 2007).

A3F has been isolated from purified HBV capsids and can inhibit the replication of HBV in the absence of hypermutation (Noguchi et al., 2005; Rosler et al., 2005; Suspene et al., 2005; Wiegand et al., 2004). However, using 3D-PCR, G-to-A hypermutated HBV genomes (context 5'-TC, minus strand) have been found in the serum of patients with chronic HBV at low frequencies (Haché et al., 2005; Suspene et al., 2005). Furthermore, A3F has been shown to inhibit both LTR retroelements such as IAP, MusD, and Ty1 (Boderd et al., 2006; Chen et al., 2006; Dutko et al., 2005; Schumacher et al., 2005). This A3 protein is also a potent inhibitor of L1 retrotransposition in the absence of hypermutation (Table II) (Muckenfuss et al., 2006; Stenglein et al., 2006).

**APOBEC3H**

Human APOBEC3H (hA3H) is expressed in the PBLs, thymus, testis, ovary, brain, small intestine, colon, and keratinocytes (Table I) (Refsland et al., 2010). This A3 protein has cytonuclear localization and is poorly expressed in vivo (Doehle et al., 2005; Hultquist et al., 2011; Kidd et al., 2007). A3H contains one cytidine deaminase catalytic domain (CDA); therefore, A3H lacks the second CDA that mediates RNA binding, homodimerization, and virion incorporation (Gooch et al., 2008; Navarro et al., 2005). However, the CDA of hA3H can inhibit the replication of both non-LTR retrotransposons such as L1 and Alu (Holmes et al., 2007; OhAinle et al., 2008). Interestingly, hA3H has been reported to hypermutate the genomes of both HBV and HPV (Table II) (Lecossier
et al., 2003). Thus far, hA3H is composed of five single nucleotide polymorphisms (SNP) (amino acid residues N15Δ, R18L, G105R, K121D, and E178D) and seven haplotypes (I-VII) (Harari et al., 2008; Wang et al., 2011). These various SNP mutations are readily found as variations in the different haplotypes. A3H mRNA can be detected; however, the protein expression of A3H has not been observed (Li et al., 2010; OhAinle et al., 2006). The reason for an absence of protein expression can be explained by a premature termination codon (PTC) in the fifth exon of the human A3H gene. If the PTC is repaired, high levels of hA3H protein are observed in cell culture (Dang et al., 2008). When the N15Δ or G105R mutations are present, there is a dramatic decrease in the expression of A3H (OhAinle et al., 2008). Only haplotype II (hapII) of hA3H, which does not have the N15Δ and/or G105R mutations, can be stably expressed. Hap II was originally identified as being the most common haplotype among the African-American population (OhAinle et al., 2008). More recently, Hap V was detected frequently in African-American, Caribbean, and Chinese populations whereas Hap VII was identified in only European Caucasian populations (Wang et al., 2011). Both hap V and hap VII have been shown to incorporate into HIV-1 virions using the \textsuperscript{112}YYXW\textsuperscript{115} motif instead of the CDA like various other A3 proteins such as hA3G. Also, both of these haplotypes inhibit HIV-1 replication. However, it appears that all the various haplotypes of human A3H are resistant to HIV-1 Vif; therefore, further research needs to be conducted to determine the antiviral activities of hA3H on HIV-1 replication (Wang et al., 2011).

Humans, chimpanzees, and the rhesus macaque (Macaca mulatta) have undergone extensive positive selection and locus expansion, but appear to share similar locus architectures. While the A3 proteins in humans have been excessively studied,
less is known about the A3 proteins in the rhesus macaque (*macaca mulata*). Rhesus A3B, C, D, F, G, and H have been identified, while rhA3A has yet to be identified (Hultquist et al., 2011; Virgen et al., 2007). RhA3D, rhA3F rhA3G, and rhA3H all contribute to the antiviral restriction of HIV-1Δvif and are targeted for degradation by SIVmac239 (Hultquist et al., 2011, Virgen et al., 2007; Zennou and Bieniasz, 2006). These homologous A3 proteins of both human and the rhesus macaque have comparable steady-state localization as well as both HIV-1 and SIV<sub>mac</sub> restriction abilities such as hypermutation in the absence of Vif (Hultquist et al., 2011; Virgen et al., 2007). However, further studies will be needed to identify rhA3A, assess the expression patterns of these proteins *in vivo*, and determine their significance in the rhesus macaque.

**Vif-APOBEC3 Paradigm**

Vif reduces the encapsidation of A3G into virions by 99% either directly inhibiting the incorporation of A3G or more importantly by depleting the cellular stores of A3G (Kao et al., 2003; Mariani et al., 2003; Marin et al., 2003; Mehle et al., 2004; Sheehy et al., 2003; Stopak et al., 2003). This ensures that the progeny virions will be highly infectious. Vif uses multiple protein-protein interactions to ensure that the antiviral effects of A3G are inhibited. The N-terminus of Vif is required for interaction with genomic RNA and to bind/interact with various regions of the A3 proteins, which will be discussed in detail below, while the C-terminus of Vif is essential to recruiting an active E3-ubiquitin ligase complex.
Sequence analysis from different lentiviral Vif proteins display several highly conserved domains in the C-terminus of Vif that are required for the assembly of the Vif-CBF-β-Cullin 5/Elongin B/Elongin C/ Nedd8/Ring box-1 E3 ubiquitin ligase complex. Vif accomplishes this by binding A3 and then a substrate recognition subunit of the Cul5/EloB/C complex targeting A3 for polyubiquitination and subsequent degradation by the 26S proteosome (Figures 10 and 11) (Conticello et al., 2003; Jäger et al., 2011; Marin et al., 2003; Sheehy et al., 2003; Stopak et al., 2003; Yu et al., 2003; Yu et al., 2004). These domains in HIV-1 Vif are the 108HCCH139 motif, 144SLQ(Y/F)LA150 domain, E88WRRKK93 and the P161PLP164 domain.

The highly conserved zinc binding domain of HIV-1 Vif is referred to as the HCCH domain and is located between amino acid residues (108-139) (Figure 12). The 108H-X5-C—X17,18-C-X3,5-H139 domain is indispensable for Vif function and thought to maintain a structural conformation of Histines and cysteines flanked by hydrophobic residues that coordinate zinc to form the Cul 5 binding site (Luo et al., 2005; Mehle et al., 2006; Xiao et al., 2006; Xiao et al., 2007). Zinc binding to the HCCH motif is both specific and reversible producing conformational changes leading to protein aggregate formation (Paul et al., 2006). These high order protein assemblies alter protein conformation to expose possible protein-protein interaction sites (Giri and Maynard, 2009; Paul et al., 2006). Upon further investigation using fluorescence and CD spectroscopy, the binding of the HCCH domain to zinc confirmed that zinc binding increases tertiary packing supporting the theory that zinc binding and protein conformation are coupled tightly (34). Chelation of zinc by cell-permeable N, N′, N″, N‴-tetrakis-(2-pyridylmethyl)ethylenediamine (TPEN) results in the loss of Vif function.
Figure 10. Proteosomal degradation of human A3G by HIV-1 Vif. 1) HIV-1 Vif binds to hA3G and recruits the E3 ubiquitin ligase complex leading to the ubiquitination and subsequent proteosomal degradation of hA3G. 2) HIV-1 Vif can impair the translation of hA3G mRNA. 3) HIV-1 Vif can physically bind hA3G excluding it from the virion and sequestering hA3G away from the sites of viral assembly. Figure adapted from Chiu and Greene, 2008.
Figure 11. Vif induced polyubiquitination of A3G through the recruitment of the CBF-β-Cul5-EloB/C-Rbx 1 E3 ubiquitin ligase complex and the specific amino acid motifs that mediate protein-protein interactions.
Figure 12. The HCCH and BC box domains of Vif are highly conserved in primate lentiviruses. The HCCH domain is indicated in orange and the BC box domain is depicted in blue.
HCCH domain

BC box

SIV_{mac} 239: STYFP CFTAGEVRR AIRGEQLLLSCCRFPRAHKYQVPSLQYLALKVVS D V R S Q Q G E N P T W K Q Q
SIVsm: STYFP CFAAHAVRQAIRGEQVLSCGYAVAHHSSVQLQYLLALKVVLQNDRPKGNPTRKK
SIVagm: PCFTDRAIQQAIRGESFLWCTYKEGHAVAENHGWQVRSLQFLALTVDTRLNRRKRFQGKA
SIVsyk: VHSQYFCFSDRAVQQALRGKLETSHCWNFHKEQVLSLQYLALQKYLSKGDGFLQSLPAA
SIVmnd: HWKYLPSCFTEQAIRQALLGKRLTVCYFHWHGSKKVGSLQYLALLSYTAYCNN GRGPRDPS
SIVcpz: HLQYFDCAFASAI RKAVLGKQVYPKCEYPAGHQVQVGLQYRALA WRVGKPPPPLPSVTK
HIV-1: HLYYFDCHSDSAIRNAILGHI V SPSCYEYQVGHNKVGSLQYLAL AALITPKRIKPPLPSVKK
HIV-2: STYFSCF TAGEVRRAIRGEKL LSCCNYPQA HKAQVPSLQYLALVVVQQNDRPQRKGTARKQ

Consensus BC box: (P,S,T)_{1}L_{2}XXX(C,A,S)_{9}XXX(A,I,L,V)_{10}
leading to an increase in antiviral activity by A3G (Xiao et al., 2007). Mutation of either cysteine residue or spacing within the HCCH motif results in the inability of Vif to target A3G for proteosomal degradation (Paul et al., 2007; Xiao et al., 2006). This is due to a conformational change in the HCCH motif that prevents the exposure of possible protein-protein interaction sites (Giri and Maynard, 2009; Paul et al., 2006). Upon further investigation using fluorescence and CD spectroscopy, the binding of the HCCH domain to zinc confirmed that zinc binding increases tertiary packing supporting the theory that zinc binding and protein conformation are coupled tightly (34). Chelation of zinc by cell-permeable \( N, N, N', N' \)-tetrakis-(2-pyridylmethyl)ethylenediamine (TPEN) results in the loss of Vif function leading to an increase in antiviral activity by A3G (Xiao et al., 2007). Mutation of either cysteine residue or spacing within the HCCH motif results in the inability of Vif to target A3G for proteosomal degradation (Paul et al., 2007; Xiao et al., 2006). Therefore, it would be interesting to see how mutations in the HCCH domain affect HIV-1 pathogenesis in vivo.

Another highly conserved domain crucial for the function of Vif is the \(^{144} \text{SLQ(Y/F)LA}^{150} \) motif which binds the cullin adapter protein EloBC through a BC-box motif (Figure 12) (Yu et al., 2003). The BC-box of the suppressors of cytokine signaling (SOCS) proteins is a loop helix motif with the consensus sequence \(((A/P/T/S)_1 \text{L}_2 \text{xxx(C/A/S)}_3 \text{xxx(A/V/L/I)}_10)\) that also exists in other cellular proteins known to interact with EloBC (Kamura et al., 1998). The serine at amino acid position 144 is regulated by phosphorylation thus negatively regulating the binding of Vif to EloC. If this serine is mutated to an alanine, phosphorylation is prevented, but this mutant still efficiently depletes cellular stores of A3G. However, the virions that are produced are
severely reduced in infectivity (Mehle et al., 2004). While Vif binds EloBC heterodimer with high affinity, it does not fit the consensus sequence perfectly. The BC-box of Vif is unstructured in its unbound state; however, this motif binds EloC through two α-helices and becomes structured. Once EloC is bound, the P^{161}PLP^{164} domain (described below) interacts with the amino acid residues 101-104 of EloB (Bergeron et al., 2010). These two binding steps are essential for the Vif-EloBC interaction and recruit the E3 ligase. If either EloB or C are deleted or mutated, the Vif-EloBC interaction is lost (Mehle et al., 2004; Yu et al., 2001; Yu et al., 2003). More studies are needed to look at the in vivo pathogenesis by making site-directed mutations in the SLQ(Y/F)LA domain of Vif.

The central hydrophobic domain E^{88}WRRKK^{93} and the conserved P^{161}PLP^{164} domain of HIV-1 Vif are important for the steady-state levels of Vif and for interactions with tyrosine kinases (Donahue et al., 2008; Fujita et al., 2003; 2004). Mutation of either of these sites results in decreased Vif-Vif multimerization, diminished Vif expression, reduced A3G binding, and a loss of viral infectivity (Fujita et al., 2003; 2004; Kataropoulou et al., 2009; Yang et al., 2001; Yang et al., 2003). Furthermore, if the PPLP motif is mutated or deleted, HIV-1 in the presence of Vif is unable to exclude A3G from the virion. This dominant negative effect leads to a decrease in virion infectivity (Walker et al., 2010). Although the role of the PPLP domain not completely clear, it has been shown that the PPLP domain may help promote the dimerization of Vif in the absence of Cul 5, that it directly interacts with Cul 5, and that it is essential for Vif function (Simon et al., 1999; Yang et al., 2001; Yang et al., 2003).
As discussed above, HIV-1 and HIV-2 were transmitted to humans by different zoonotic transmission events (Sharp et al., 2001). Briefly, SIVcpz (chimpanzee, *Pan troglodytes*) was the precursor for HIV-1 and SIVsm (sooty mangabey, *Cerocebus*) was the source of HIV-2, which are both related but very distinct SIVs that naturally infect nonhuman primates (Sharp et al., 2001). These zoonotic infections seem to occur when the host cell does not strongly resist lentiviral infection with post-entry restriction factors, such as TRIM5α or A3G (Cullen et al., 2006; Stremlau et al., 2004). The Vif-A3G interaction is an important factor in species-specificity between HIV and various SIVs. Zoonotic infections only persist when the lentiviral Vif protein can bind to and induce the degradation of A3G (Cullen et al., 2006; Mariani et al., 2003). For example, HIV-1 and SIVcpz Vif can bind and degrade human A3G but not African green monkey A3G rendering humans resistant to infection by SIVagm. On the contrary, SIVagm cannot bind and degrade HIV-1 or SIVcpz Vif (Cullen et al., 2006; Mariani et al., 2003). The direct interaction between A3G and Vif are dependent on amino acids 128-130 in A3G (Huthoff and Malim, 2007; Russell et al., 2009). Previous research has determined that species-specificity is determined by a single amino acid in human A3G, D128, and Africa green monkey A3G, K128 (Bogerd et al., 2004; Mangeat et al., 2004; Schrofelbauer et al., 2004; Xu et al., 2004) and amino acids 14-17 on HIV-1 Vif (Schrofelbauer et al., 2006). In human A3G if the aspartic acid is replaced with a lysine, hA3G becomes sensitive to SIVagm Vif and insensitive to HIV-1 Vif (Bogerd et al., 2004; Mangeat et al., 2004; Schrofelbauer et al., 2004; Xu et al., 2004). Also, if the D\textsuperscript{14}RMR\textsuperscript{17} motif in HIV-1 Vif is replaced with either the S\textsuperscript{14}ERQ\textsuperscript{17} or S\textsuperscript{14}EMQ\textsuperscript{17} residues...
in SIVagm Vif, the HIV-1 Vif protein is now able to interact with both A3G from African
green monkeys and rhesus macaques (Schrofelbauer et al., 2006). These species-
specific limitations of Vif form a means to prevent or minimize the number of successful
zoonotic transmissions of various primate lentiviruses.

Regions implicated ability of Vif to bind and interact with A3 proteins have been
mapped to the N-terminal region of HIV-1 Vif (Figure 13). The amino acid residues
involved are arranged in a non-linear fashion thereby indicating that multiple surfaces
are involved in this interaction. The exact binding domain in HIV-1 Vif consists of
several discontinuous subdomains (Goila-Gaur et al., 2008, Marin et al., 2003; Mehle et
al., 2007; Russell et al., 2007; Russell et al., 2009; Schrofelbauer et al., 2006; Simon et
al., 2005; Tian et al., 2006). For example, using site-directed mutagenesis, hydrophilic
motifs \(^{40}YRHHY^{44}\) and \(^{23}SLVK^{26}\) were found to be essential for the restriction of A3G or
A3G and A3F, respectively (Chen et al., 2009; Dang et al., 2009; Marin et al., 2003;
Russell et al., 2007; Russell et al., 2009; Yamashita et al., 2008; Zhang et al., 2008)
The hydrophobic patch, \(^{69}YWxL^{72}\) was shown to suppress A3G, A3F, and A3B and the
\(^{55}VxIPLxL^{64}\) motif is important for binding A3G (Walker 18, 50; He et al., 2008; Pery et al.,
2009). These residues 55-72 are sufficient alone for A3G binding and the charged
residues in this region (R56, R61, and R63) are dispensable for A3G degradation (He et
al., 2008; Mehle et al., 2007). The \(^{12}QVDRMR^{17}\) motif is important for interaction with
A3F and A3C (Mehle et al., 2007; Russell et al., 2007). The \(^{81}LGxGxxIxW^{89}\) domain
has been found to bind and regulate both A3G and A3F neutralizing activities. The L81
and G82 regulates the binding of Vif to A3F, while G84, I87, and W89 regulate
interaction with both A3G and A3F (Dang et al., 2010). Both the \(^{74}TGERxW^{79}\) and
EDRWN$^{175}$ domains specifically bind and neutralize A3F (Dang et al., 2010; Russell et al., 2007).

Although there is only approximately 50% homology between HIV-1 Vif and SIV$_{mac239}$, very few domains in the N-terminus of HIV-1 Vif are partially conserved in SIV$_{mac239}$ Vif (Dang et al., 2009; He et al., 2008; Pery et al., 2009). These domains are listed in Figure 13. Identifying these domains on SIVmac239 Vif is essential for using the rhesus macaque as a model to study viral pathogenesis. Even though the interaction surfaces on Vif are structurally complicated, it is important to understand and identify these domains to determine the mechanism by which Vif recognizes A3 proteins.

**APOBEC3 and Vif as Therapeutic Targets**

Since a protective vaccine is not available, the Merck vaccine trial proved unsuccessful, and the resistance to HARRT therapy occurs, we need new and more effective antiviral drugs for the management of HIV-1 infection (Barouch et al., 2008; Robb et al., 2008). Some scientists are now looking into using the Vif-A3 protein interactions as targets for antiviral drug development. This is possible because several studies have implied that interference with the Vif-A3 interaction strongly suppress HIV-1 viral replication in nonpermissive cells (Gabuzda et al., 1992; Fisher et al., 1987; Strebel et al., 1987; von Schwedler et al., 1993).

In order to use APOBEC3 proteins as therapeutics, the significant quantities of APOBEC3 needs to be incorporated/packaged into the HIV-1 virion. Using this
Figure 13. The domains of HIV-1 and SIV\textsubscript{mac239} Vif that have been implicated in binding A3 proteins. Shown are representations of both the N-terminus of HIV-1 (top) and SIV\textsubscript{mac239} (bottom) Vif from amino acids 1 to 174 as indicated below the protein. Labeled in the boxes are specific amino acid motifs that are required to interact with the designated A3 proteins listed above the proteins.
scenario, even HIV-1 virions would contain increased amounts of A3 effectively overwhelming the normal Vif function thus preventing the genomic RNA from becoming a replication-competent provirus in the target cell (Albin et al., 2010; Jern et al., 2009; Mulder et al., 2008). The amount of risk associated with increasing A3G expression and causing mutagenic activity would need to be smaller than the ability of A3G to exert its antiviral activities. In order to increase encapsidation of A3G into HIV-1 virions and prevent the degradation of A3G, two different approaches can be employed: 1.) Directly increasing incorporation of A3G by increasing the expression of A3G, or 2.) Indirectly increasing A3G incorporation by preventing proteosomal degradation of A3G (Albin et al., 2010).

In order to directly increase the expression of A3G thus increasing the incorporation, IFN-α treatment is possible. Induction of A3 proteins in macrophages by IFN-α treatment is effective in restricting the infectivity of HIV-1 (Peng et al., 2006). By doing this, we may be able to selectively increase A3G expression and aid the immune system in controlling HIV-1 infection. This treatment could also enhance other small molecules that occur naturally to aid the immune system or inhibit viral replication. However, IFN-α treatment induces many non-specific cellular effects making it difficult to rule out the A3 antiviral activity. Another side effect of this treatment could also be the hypermutation of the human genome because A3 expression levels would overwhelm the cellular regulatory mechanisms leading to the development of certain cancers (Harris et al., 2002; Yamanaka et al., 1995).

Another direct approach could also be to selectively target A3 proteins to the viral core in HIV-1 virions. For example; if A3G is fused to a fragment of Vpr, A3G is
efficiently incorporated into HIV-1 virions in the presence of Vif (Aguiar et al., 2008). Also, fusing A3A to Vpr resulted in enhanced antiviral activity in the target cell. This is important because normally A3A does not restrict HIV-1 or reach the viral core; however, when fused to Vpr, the A3A-Vpr protein is specifically targeted to the viral core and potently restricts HIV-1 (Aguiar et al., 2008).

An indirect approach would be to increase incorporation of the A3 protein and to prevent its proteosomal degradation. For example, in one study a membrane-permeable zinc chelator, TPEN was used to prevent the formation of the E3 ubiquitin ligase by inhibiting the Vif-Cullin 5 binding thus preventing the degradation of A3G (Xiao et al., 2007). However, TPEN in vitro has been shown to induce apoptosis in human cells by inhibiting other cellular functions, which would make this chelator undesirable (Martin et al., 1991; Cao et al., 2005). Also, proline-rich peptides containing the PPLP motif have been used to prevent the oligomerization of Vif in nonpermissive cells and inhibit the replication of HIV-1 (Miller et al., 2007; Yang et al., 2003). Peptide inhibitors, such as these, provide useful information on protein-protein interactions, but are not desirable as antiviral drugs due to their low stability and poor bioavailability (Adessi et al., 2002).

In order to indirectly increase incorporation of A3 proteins into HIV-1 virions, the interacting domains of both Vif and A3 need to be elucidated. Currently, in vitro Vif-A3G binding assays are underway using a fluorescence resonance energy transfer (FRET) assay to measure the interactions between GST-Vif and a biotinylated A3G peptide. High-throughput screening is being used to identify potential candidate molecules that may target the Vif-A3G interaction (Mehle et al., 2007).
Another indirect approach may be to create a molecule that directly binds A3 and inhibits the binding of HIV-1 Vif (Harris et al., 2004). Also, it could be theoretically possible to create a molecule that could be used to inhibit the interaction of HIV-1 Vif with either A3 proteins or the proteosomal machinery. For example, RN-18 (containing a 2-(4-nitophenylthio-N-phenylbenzamide) moiety) is a small-molecule antagonist of HIV-1 Vif that inhibits the interaction between A3-Vif (Nathans et al., 2008). RN-18 increases the expression levels of A3G, A3F, A3C as well as the incorporation of A3G into the virion, but it does not affect the binding of HIV-1 Vif to A3G or A3F. The inhibitor does increase the proteosomal degradation of Vif in the presence of A3G, A3F, or A3C. Also, RN-18 did not display cytotoxic effects at a concentration of 50-100 µM (Nathans et al., 2008). However, this inhibitor is a recent development and is not yet ready for clinical trials. While there has been promising success in designing small molecule inhibitors, more research is needed to visualize the structural interaction between A3 and Vif to design efficient interaction inhibitors.

The final approach may be the use of gene therapy. However, further advances in both safety and efficacy of both cell therapy and gene delivery must be overcome. As mentioned above, the Vif-A3 interaction is species-specific. If human A3G aspartate 128 is substituted with the lysine found in both the rhesus macaque and the African green monkey A3G, the human A3G would become resistant to Vif-mediated proteosomal degradation (Mariani et al., 2003). The human D128K A3G mutation would protect hA3G from binding HIV-1 Vif (Bogerd et al., 2004; Huthoff et al., 2007; Li et al., 2008; Mangeat et al., 2004; Schröfelbauer et al., 2004). Theoretically this would lead to increased packaging of A3G into the HIV-1 virion by preventing its interaction
with Vif thus leading to the restriction of HIV-1. To date, there are no known occurrences of this allele circulating in the human population; therefore, therapeutic gene delivery would be necessary. Another method would be to fuse A3G to ubiquitin-associated domain 2 (UBA2), a stabilization signal known to protect proteins from proteosomal degradation (Li et al., 2008). This fusion only has a partial effect on restriction of HIV-1 and is only partially restricted to Vif-mediated degradation (Li et al., 2008). This modification along with others may someday provide a tactic in engineering T cells to be resistant to HIV-1 infection.

**Goals of the study**

All primate lentiviruses encode for a Vif protein, which is necessary for HIV-1 replication in primary CD4+ T cells and macrophages (Fan and Peden, 1992; Gabuzda et al., 1992; Blanc et al., 1993; Sakai et al., 1993; von Schwedler et al., 1993; Borman et al., 1995). The Vif protein interacts with apolipoprotein B mRNA-editing enzyme catalytic peptide-like 3 (APOBEC3; A3) proteins promoting their accelerated degradation by the 26S proteosome (Sheehy et al., 2002). Sequence analysis of Vif proteins from different lentiviruses revealed that there are two highly conserved domains in the carboxyl terminus that are required for the recruitment of the Vif-CBF-β-Cul5/Elongin B/C/Rbx-1 E3 ubiquitin ligase complex (Hultquist et al., 2012; Luo et al., 2005; Mehle et al., 2004; 2006; Yu et al., 2004). These domains are the viral BC box, SLQ(Y/F)LA, and the Zn2+ (H-X5-C-X17-18-C-X3-5-H; HCCH) binding domains. Previous cell culture studies have shown that the introduction of amino acid substitutions in the SLQ(Y/F)LA motif resulted in decreased binding of Vif to Elongin C, while substitutions
in the HCCH domain prevented interactions with Cullin 5 in the Vif-CBF-β-Cul5/Elongin B/C/Rbx-1 E3 ubiquitin ligase complex (Luo et al., 2005; Mehle et al., 2004a; 2004b; 2006; Stopak et al., 2003; Yu et al., 2203; 2004).

In our first study we were interested in determining the importance of the highly conserved SLQYLA motif in viral pathogenesis using the SHIV/macaque model of infection. Two amino acid substitutions were introduced into the SLQYLA (AAQYLA) domain of the SIV Vif protein. Prior to our rhesus macaque study, the resulting virus, SHIV<sub>VifAAQYLA</sub>, was assessed in both non-permissive (CEM) and permissive (CEM-SS) cells to examine the replication of these mutant viruses relative to parental SHIV. We found that, while SHIV<sub>VifAAQYLA</sub> replicated with similar kinetics to SHIV<sub>KU-1bMC33</sub> in CEM-SS cells, it did not replicate in CEM cells. We were also interested in determining if human A3G is incorporated into SHIV<sub>VifAAQYLA</sub> and showed that human A3G was readily incorporated into SHIV<sub>VifAAQYLA</sub> virions but not SHIV<sub>KU-1bMC33</sub>. To assess the replication and persistence of SHIV<sub>VifAAQYLA</sub> <em>in vivo</em>, SHIV<sub>VifAAQYLA</sub> was inoculated into three rhesus macaques where various viral and immunological parameters were followed for six months. All three macaques inoculated with SHIV<sub>VifAAQYLA</sub> developed a transient decrease in circulating CD4+ T cells one week post-inoculation that rebounded to pre-inoculation levels by four weeks post-inoculation. These macaques also had plasma viral loads that were 100-fold lower than macaques inoculated with SHIV<sub>KU-1bMC33</sub>, developed no histological lesions in lymphoid tissues, and had a decreased distribution of replicating virus in visceral tissues. The engineered Vif mutations were stable during the first three weeks post-inoculation, but by four weeks post-inoculation the S147A amino acid substitution changed to a threonine in two of the three macaques. At
necropsy, plasma was analyzed for the presence of immunoprecipitating antibodies, and all three macaques developed immunoprecipitating antibodies although one macaque developed significantly less. Sequence analysis of vif, vpu, env, and nef from the thymus, mesenteric, axillary, and inguinal lymph nodes, revealed an increased amount of G-to-A substitutions in the dinucleotide context of 5'-TC (minus strand) in these genes amplified from macaques inoculated with SHIV_{VifAAQYLA}. To determine if pathogenic virus was present in plasma from macaques inoculated with SHIV_{VifAAQYLA} at the time of necropsy, plasma was pooled from two of the three macaques and inoculated into a naïve macaque (macaque I95). Macaque I95 was assessed for various viral parameters such as circulating CD4^{+} T cell levels, virus burden, immunoprecipitating antibodies, and the detection of replicating virus in both PBMC and visceral tissues. We showed that this macaque maintained circulating CD4^{+} T cell levels, no viral loads, and no immunoprecipitating antibodies. Viral DNA sequences (the gag gene) could be detected in PBMC through out the duration of infection; however, viral RNA sequences could not. At necropsy, viral DNA and RNA could be detected in the liver, kidney, spleen, tonsil, axillary lymph node, and small intestine (ileum). This study indicated that a certain level of infectious viral replication was occurring in macaques inoculated with SHIV_{VifAAQYLA} and that the pooled plasma contained only low levels of infectious virus that was efficiently controlled by macaque I95.

In our second study we used the SHIV/macaque model of infection to compare the replication and pathogenicity of SHIVs that express a Vif protein in which the entire SLQYLA (SHIV_{Vif5A}) or HCCH (SHIV_{VifHCCH(-)}) domains were substituted with alanine
residues. We were interested in extending our previous studies to determine if these highly conserved domains were critical to Vif function in vivo. Prior to our rhesus macaque study, we were interested in determining whether SHIV_{Vif5A} and SHIV_{VifHCCH(-)} were effectively inhibited by rhesus A3G/F. We performed assays to examine the replication of these mutants in both non-permissive (C8166, PBMC) and permissive (SupT1) cells. We also determined if SHIV_{Vif5A} and SHIV_{VifHCCH(-)} in the presence of rhesus A3G/F incorporated into virions, remained stably expressed in cells, restricted virus replication, and if these viruses were capable of introducing significant G-to-A substitutions into the nef gene of the SHIV genome in vitro. We found that both SHIV_{Vif5A} and SHIV_{VifHCCH(-)} did not replicate in non-permissive cells but replicated efficiently in permissive cells. We also showed that rhesus A3G/F was readily incorporated into SHIV_{Vif5A} and SHIV_{VifHCCH(-)} virions, remained stably expressed in the cell, severely restricted viral replication, and induced a significant number of G-to-A substitutions in nef similar to SHIV_{VifStop} (SHIV expressing a Vif protein with two stop codons introduced at amino acid residues 28 and 29). To assess the replication and persistence of these viruses in vivo, each virus was inoculated into three rhesus macaques where various viral and immunological parameters were followed for six months. All macaques maintained a stable level of circulating CD4+ T cells, low viral burdens, maintained engineered mutations, yielded no histological lesions, and developed immunoprecipitating antibodies by 12 weeks post-inoculation. However, the production of viral RNA only persisted in macaques inoculated with SHIV_{VifHCCH(-)}. Sequence analysis of nef and vpu from the small intestine (ileum), thymus, and the spleen showed extensive G-to-A substitutions in nef genes isolated from macaques.
inoculated with SHIV_{VifHCCH(-)}. Macaques inoculated with SHIV_{Vif5A} effectively controlled the virus three weeks post-inoculation and no viral sequences could be amplified from tissue DNA. These studies showed that the SLQYLA and HCCH domains are critical for viral pathogenesis in vivo and that there may exist APOBEC3 negative reservoirs in the rhesus macaque that allow for low levels of viral replication and persistence but not disease. Therefore, from these first two studies we inferred that mutations targeted to one or more functional conserved domains within the Vif protein may limit viral replication and generate an effective immune response leading to the “self-inactivation” of the virus by the activities of various APOBEC3 proteins resulting in a possible live-attenuated vaccine candidate.

While the human APOBEC3 repertoire has been extensively studied, the full complement of these proteins in the rhesus macaque remains unknown. Sequencing of the rhesus macaque genome has led to the identification of the rhesus homologues A3B, A3C, A3D, A3F, A3G, and A3H. In the final study, we were interested in identifying and characterizing the human A3A (hA3A) homologue in the rhesus macaque (rhA3A). Using cDNA from Concanavalin-A-activated rhesus macaque PBMCs, we were able to detect and clone rhesus A3A. Using the cloned rhesus A3A sequence we were able to determine that both the human and rhesus Apobec3 genes are orthologous. We show that rhA3A is highly expressed in activated CD4^+ T cells, widely expressed in both the visceral and central nervous system tissues of the rhesus macaque, maintains nucleocytoplasmic localization, does not inhibit virus release, does not restrict foreign DNA, and is degraded in the presence of the human immunodeficiency virus (HIV-1) and simian-human immunodeficiency virus (SHIV).
genomes in a Vif-dependent manner. Our results also indicate that rhA3A reduced the level of infectious SHIVΔvif by approximately 20-fold and HIV-1Δvif by 3-fold. Human and monkey A3A amino acid sequences are 81% homologous and can be distinguished by a three amino acid indel (insertion and/or deletion) located between residues 27-30. When these residues were deleted from rhA3A (rhA3AΔSVR), the antiviral activity of rhA3A was abolished suggesting that these residues are critical for lentivirus inhibition. Select APOBEC3 proteins are incorporated into the virion and can inhibit reverse transcription and/or induce G-to-A hypermutation in nascent reverse transcripts in the next target cell. Previous studies revealed that rhA3G is incorporated into SHIVΔvif virions and exerts its antiviral activity in target cells by an increase in cytidine deamination of newly synthesized minus-strand viral DNA from cytosines to uracils, leading to G-to-A substitutions (plus strand) in the viral genome. We were able to detect the incorporation of rhA3A into SHIVΔvif and to a lesser extent in SHIV virions; however, we were unable to detect the incorporation of rhA3A into either HIV-1 or HIV-1Δvif virions. Even though rhA3A is incorporated into SHIVΔvif virions and potently restricts SHIVΔvif similar to rhA3G, rhA3A produced an approximately 5-fold decrease in the number of G-to-A mutations compared to rhA3G. Since previous research showed that human A3A inhibits the replication of adeno-associated virus 2 (AAV-2), intracisternal A particles (IAP), and long interspersed nuclear element 1 (L1), we were interested in determining if rhesus A3A could inhibit AAV-2 replication and L1 retrotransposition (Bogerd et al., 2006a; 2006b; Chen et al., 2006; Muckenfuss et al., 2006). We found that rhesus A3A could not inhibit AAV-2 replication or L1 retrotransposition. The data from this study suggests for the first time that a primate
A3A protein can inhibit lentiviral replication but not AAV-2 replication or L1 retrotransposition.
VI. Chapter Two: Mutations in the SLQYLA Motif of Vif in the Simian-Human Virus Results in a Less Pathogenic Virus in Rhesus Macaques

Abstract

The simian-human immunodeficiency virus (SHIV)/macaque model for human immunodeficiency virus type 1 (HIV-1) has become a useful tool to assess the role of accessory genes like virion infectivity factor (Vif), in lentiviral pathogenesis. In this study, we introduced two amino acid changes in the highly conserved SLQYLA domain (to AAQYLA) of the SIV Vif protein. The resulting virus SHIV\textsubscript{VifAAQYLA}, was used to infect three macaques, which were followed for over six months. Plasma viral loads and circulating CD4\textsuperscript{+} T cell levels were assessed during the course of infection. The three macaques inoculated with SHIV\textsubscript{VifAAQYLA} did not develop significant CD4\textsuperscript{+} T cell loss over the course of their infection, had plasma viral loads that were over 100-fold lower than macaques inoculated with parental SHIV\textsubscript{KU-1bMC33}, and developed no histological lesions in lymphoid tissues. DNA and RT-PCR analysis revealed that only a select number of tissues were infected with this virus. Sequence analysis indicates that the site-directed changes were stable during the first three weeks after inoculation but thereafter the S147A amino acid substitution changed to a threonine in two of the three macaques. However, the L148A substitution remained stable in the \textit{vif} amplified from the PBMC of all three macaques. Sequence analysis of \textit{vif}, \textit{vpu}, \textit{env}, and \textit{nef} genes revealed G-to-A substitutions in the genes amplified from macaques inoculated with SHIV\textsubscript{VifAAQYLA}, which were higher than in a macaque inoculated with parental SHIV\textsubscript{KU-1bMC33}. We found that the majority (>85%) of the G-to-A substitutions were in the context of 5'-TC (minus
strand) and not 5′-CC, suggestive that one or more the rhesus APOBEC3 proteins may be responsible for the observed mutational patterns. Finally, macaques inoculated with SHIV\textsubscript{VifAAQYLA} developed immunoprecipitating antibody responses against the virus. The results from this study provide the first \textit{in vivo} evidence of the importance of the conserved SLQYLA domain in viral pathogenesis and show that targeted mutations in \textit{vif} can lead to a persistent infection with G-to-A (plus strand) changes accumulating in the viral genome.

\section*{Introduction}

Human immunodeficiency virus type 1 (HIV-1) as well as other lentiviruses encode for a Vif protein, which has been shown to be essential for HIV-1 replication in certain cell types. The Vif protein of HIV-1 was first shown to interact with apolipoprotein B mRNA editing enzyme catalytic polypeptide-like 3G (APOBEC3G; hA3G) \citep{Sheehy2002}. This protein was found to be incorporated into the virus and provide cells with an innate intracellular anti-retroviral activity that is associated with hypermutation of the viral genome through cytidine deamination \citep{Harris2003, Lecossier2003, Mangeat2003, Zhang2003}. Several groups subsequently showed that the Vif protein can prevent hypermutation by binding to hA3G and targeting this protein for degradation via the proteasome \citep{Conticello2003, Kao2004, Mariani2003, Marin2004, Zhang2003}. This hA3G-induced cytidine deamination results in cytidine to uridine changes during minus strand DNA synthesis triggering the DNA repair pathways that lead to the degradation of viral transcripts \citep{Harris2003, Mangeat2003, Mariani2003, Yu2004, Zhang2003}.
al., 2003; Mehle et al., 2004; Sheehy et al., 2003; Stopak et al., 2003; Yu et al., 2003; 2004). In addition to A3G, humans have six other A3 genes: hA3A, hA3B, hA3C, hA3DE, hA3F, and hA3H (Jarmuz et al., 2002). The members of this family either have one (hA3A, hA3C, and hA3H) or two (hA3B, hA3DE, hA3F and hA3G) Zn$^{+2}$ coordinating deaminase domains organized as H-x$^{-1}$-E-x$_{25-31}$-C-x$_{2,4}$-C (with x being a non-conserved position) (Chiu and Greene, 2008). Other APOBEC3 family members such as hA3B, hA3DE, hA3F, and hA3H have also been shown to also inhibit the replication of HIV-1Δvif (Dang et al., 2007; Doehle et al., 2005; Wiegand et al., 2004; Yang et al., 2007; Yu et al., 2004; Zheng et al., 2004). Moreover, SIV$^{mac}_{239}$Δvif is potently restricted by hA3G, hA3F, and hA3H and to a lesser extent by hA3B, hA3C, and hA3DE (Dang et al., 2007; 2008; Mariani et al., 2003; Virgen et al., 2007; Yu et al., 2004). While fewer studies have been performed on the APOBEC3 family proteins from macaques, two studies have shown that rhesus macaque A3G and A3F inhibit the replication of SIV$^{mac}_{Δvif}$ (Zennou et al., 2006; Virgen et al., 2007).

As mentioned, previous studies have shown that Vif inhibits the antiviral activity of APOBEC3 proteins by targeting A3 proteins for proteosomal degradation through a CBF-β-Vif-Cul5/ElonginB/C/Rbx-1 E3-dependent ubiquitin ligase (Jäger et al., 2011; Kobayashi et al., 2005; Mehle et al., 2004; Shao et al., 2010; Stanley et al., 2008; Yu et al., 2004). Vif binds the Cullin 5 complex through a highly conserved BC-box region known as the $^{144}$SLQ(Y/F)LA$^{149}$ domain (Yu et al., 2003). The BC-box of Vif is a loop-helix motif with a consensus sequence ((A/P/T/S)$_1$L$_2$x xx(C/A/S)$_6$x xx(A/V/L/I)$_{10}$) that has also been identified in other cellular proteins that are known to interact with elongin B/C (Kamura et al., 2004; Luo et al., 2005). Through this motif, Vif has a high affinity for
elongin B/C (Mehle et al., 2004). Mutation or deletion of the SLQ(Y/F)LA motif disrupts the proteosomal targeting and degradation of A3 proteins, and prevents the production of infectious HIV-1 virions (Kobayashi et al., 2005; Mehle et al., 2004a; 2004b; Shao et al., 2010; Simon et al., 1999; Stanley et al., 2008; Yu et al., 2004). While this domain has been extensively studied in vitro, no studies have assessed the role of this domain using a non-human primate model of HIV-1 pathogenesis.

In this study, we constructed a simian-human immunodeficiency virus (SHIV) with amino acid changes in the highly conserved SLQYLA (SHIV\textsubscript{VifAAQYLA}) domain and assessed it for pathogenesis in macaques. Mutation of this motif resulted in lower viral loads, stabilization of circulating CD4\textsuperscript{+} T cell levels, immunoprecipitating antibodies, and the accumulation of G-to-A substitutions in secondary lymphoid tissues. Our results show for the first time that the SLQYLA domain has determinants that contribute to the pathogenicity of SHIV in macaques and that mutations in this particular domain result in the accumulation of G-to-A mutations in the viral genome.

Results

Replication of SHIV\textsubscript{VifAAQYLA} in APOBEC3G (A3G) positive and negative cell lines

We performed assays to examine the replication of parental SHIV\textsubscript{KU-1bMC33} and SHIV\textsubscript{VifAAQYLA} in A3G/F positive (CEM) and negative (CEM-SS) cell lines. Cells were infected with each of the two viruses and the levels of p27 Gag released into the culture medium were quantified using a commercial antigen capture assay. SHIV\textsubscript{VifAAQYLA} replication in CEM-SS cells with similar kinetics to SHIV\textsubscript{KU-1bMC33} (Figure 14A).
However, SHIV\textsubscript{VifAAQYLA} released less than 0.5% p27 compared to the parental SHIV\textsubscript{KU-1bMC33} in CEM cells (Figure 14B).

**Human APOBEC3G is incorporated into SHIV\textsubscript{VifAAQYLA} virions**

Since SHIV\textsubscript{VifAAQYLA} released significantly less p27 than SHIV\textsubscript{KU-1bMC33} in CEM (A3G/F positive) cells, we determined if human A3G was incorporated into maturing virus particles leading to the restriction of SHIV\textsubscript{VifAAQYLA} replication. 293 cells were transfected with HA-hA3G and the complete genomes of either SHIV\textsubscript{VifAAQYLA} or SHIV\textsubscript{KU-1bMC33}. At 48 hours post-transfection, the culture medium was collected, clarified, and the virus partially purified and concentrated by ultracentrifugation. The amount of p27 was determined and a western blot was run to detect for the presence or absence of hA3G. The results shown in Figure 15 indicate that hA3G was incorporated into SHIV\textsubscript{VifAAQYLA} virus particles but excluded from SHIV\textsubscript{KU-1bMC33}.

**Disease in macaques inoculated with SHIV\textsubscript{VifAAQYLA}**

We inoculated three macaques with SHIV\textsubscript{VifAAQYLA} (RAK10, RCS10, and RPL10). These macaques developed a transient decrease in the level of circulating CD4\textsuperscript{+} T cells at one week post-inoculation (Figure 16B). However, by four weeks post-inoculation the level of circulating CD4\textsuperscript{+} T cells rebounded to near pre-inoculation levels (Figure 16B). These macaques maintained high levels of circulating CD4\textsuperscript{+} T cells throughout the course of their infection. All three macaques were euthanized at 28 weeks post-inoculation. At necropsy, macaques inoculated with
Figure 14. Replication of SHIV$_{KU-1bMC33}$ and SHIV$_{VifAAQYLA}$ in APOBEC3G positive (CEM) and negative (CEM-SS) cell lines. CEM and CEM-SS cells were inoculated with equal amounts of each virus and levels of p27 in the culture supernatants determined at various time points post-inoculation. (Panel A) Replication of SHIV$_{KU-1bMC33}$ and SHIV$_{VifAAQYLA}$ in CEM-SS cells. (Panel B) SHIV$_{KU-1bMC33}$ and SHIV$_{VifAAQYLA}$ in CEM cells. (○) SHIV$_{KU-1bMC33}$; (■) SHIV$_{VifAAQYLA}$. 
A. CEM-SS cells

B. CEM cells

Days post-inoculation

SHIV\textsubscript{KU-16MC33}  
SHIV\textsubscript{VsvAAGYLA}  
p27 released (ng/ml)

Days post-inoculation
Figure 15. Human APOBEC3G is incorporated into SHIV_{VifAAQYLA} virions. 293 cells were transfected with a plasmid that expresses the genome of either SHIV_{KU-1bMC33} or SHIV_{VifAAQYLA} and a vector expressing HA-hA3G. At 48 hours, the culture medium was collected, clarified, and concentrated by ultracentrifugation through a 20/60% sucrose (w/v) gradient. Equivalent levels of p27 from each sample was resuspended in 2x sample reducing buffer, boiled, and proteins separated by SDS-PAGE. The presence of hA3G was detected by Western blot using an antibody directed against the HA-tag. Lane 1. Human A3G protein in pelleted supernatant from cells transfected with empty vectors. Lane 2-4. Human A3G detected in virus from cells transfected with the vector expressing hA3G and SHIV_{KU-1bMC33} (Lane 2), SHIV_{VifAAQYLA} (Lane 3), or the positive control SHIVΔVif (Lane 4).
Figure 16. Circulating CD4$^+$ T cell levels in macaques inoculated with SHIV$_{KU-1bMC33}$ and SHIV$_{VifAAQYLA}$. (Panel A) The levels of circulating CD4$^+$ T cells in five macaques inoculated with SHIV$_{KU-1bMC33}$ (2000, ●; CM4G, ■; CM4K, ▲; RRH10, ▼). (Panel B) The levels of circulating CD4$^+$ T cells in three macaques inoculated with SHIV$_{VifAAQYLA}$ (RAK10, ●; RCS10, ♦; RPL10, ▲).
SHIV_{\text{VifAAQYLA}} had circulating CD4$^+$ T cell counts of 2427, 1656, and 3021, respectively. This contrasts with macaques inoculated with SHIV_{\text{KU-1bMC33}} (Figure 16A). Analysis of plasma viral loads in macaques inoculated with SHIV_{\text{VifAAQYLA}} revealed that during the early peak (1 to 3 weeks) of viremia, the mean load was $8.52 \times 10^4$ copies per ml (Figure 17A), approximately 100-fold less than macaques inoculated with parental SHIV_{\text{KU-1bMC33}} (Figure 17B). Following the first month of infection, the plasma viral loads in macaque RPL10 declined faster than macaques RAK10 and RCS10 (Figure 17A).

**Stability of the vif gene mutations during the course of infection and at necropsy**

We assessed the stability of the engineered mutations in the vif gene during the course of infection. DNA was extracted from PBMC samples at various times post-inoculation and from lymphoid tissues at necropsy. The vif sequences were amplified, directly sequenced and compared to the input vif sequence of SHIV_{\text{VifAAQYLA}}. The vif mutations were stable from one to three weeks post-inoculation for all three SHIV_{\text{VifAAQYLA}} infected macaques (Figure 18). However, four weeks post-inoculation the S147A amino acid substitution changed to a threonine in two of the macaques, RPL10 and RAK10. Interestingly, this amino acid substitution was mediated by a G-to-A mutation. DNA from lymphoid organs obtained at necropsy from RAK10 and RPL10 also showed the A147T mutation whereas macaque RCS10 maintained the S147A amino acid substitution (Figure 18). The L148A substitution was found to be stable during the course of infection for all three macaques and in the lymphoid organs analyzed at necropsy. There were no mutations in this region of vif from RRH10, which was inoculated with SHIV_{\text{KU-1bMC33}} (Figure 18).
Figure 17. Plasma viral loads in macaques inoculated with SHIV\textsubscript{VifAAQYLA} and SHIV\textsubscript{KU-1bMC33}. (Panel A) Plasma viral loads in three macaques inoculated with SHIV\textsubscript{VifAAQYLA} (RAK10, ●; RCS10, ◆; RPL10, ▲). (Panel B) Plasma viral loads in four macaques (2000, ●; CEM4G, ■; CEM4K, ▲; and RRH10, ▼) following inoculation with SHIV\textsubscript{KU-1bMC33}. 
Figure 18. Analysis of the vif gene for the stability of the engineered mutations. DNA was isolated from PBMC at 1, 3, 4, 8, 12, and 28 weeks post-inoculation and from lymphoid tissues (thymus, mesenteric lymph node and inguinal lymph node) at necropsy. The vif gene was amplified and directly sequence. Shown at the top are the sequences for both parental SHIV\textsubscript{KU-1\textsubscript{bMC33}} and SHIV\textsubscript{VifAAQYLA}. Amino acids depicted in red indicate the A147T mutation and the amino acids shown in blue indicate the maintained S147A amino acid substitution.
<table>
<thead>
<tr>
<th>SHIV&lt;sub&gt;KU-1bMC33&lt;/sub&gt;</th>
<th>SHIV&lt;sub&gt;VifAAQYLA&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>147-AGCCTACAGTACTTAGCA – 152</td>
<td></td>
</tr>
<tr>
<td>A A Q Y L A</td>
<td></td>
</tr>
</tbody>
</table>

**Macaque RAK10**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBMC/week 1</td>
<td>A A Q Y L A</td>
</tr>
<tr>
<td>PBMC/week 3</td>
<td>A A Q Y L A</td>
</tr>
<tr>
<td>PBMC/week 4</td>
<td>T A Q Y L A</td>
</tr>
<tr>
<td>PBMC/week 8</td>
<td>T A Q Y L A</td>
</tr>
<tr>
<td>PBMC/week 12</td>
<td>T A Q Y L A</td>
</tr>
<tr>
<td>PBMC/Necropsy</td>
<td>T A Q Y L A</td>
</tr>
<tr>
<td>Thymus</td>
<td>T A Q Y L A</td>
</tr>
<tr>
<td>Mesenteric LN</td>
<td>A A Q Y L A</td>
</tr>
<tr>
<td>Ing. LN</td>
<td>T A Q Y L A</td>
</tr>
</tbody>
</table>

**Macaque RCS10**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBMC/week 1</td>
<td>A A Q Y L A</td>
</tr>
<tr>
<td>PBMC/week 3</td>
<td>A A Q Y L A</td>
</tr>
<tr>
<td>PBMC/week 4</td>
<td>A A Q Y L A</td>
</tr>
<tr>
<td>PBMC/week 8</td>
<td>A A Q Y L A</td>
</tr>
<tr>
<td>PBMC/week 12</td>
<td>A A Q Y L A</td>
</tr>
<tr>
<td>Thymus</td>
<td>A A Q Y L A</td>
</tr>
<tr>
<td>Mesenteric LN</td>
<td>A A Q Y L A</td>
</tr>
<tr>
<td>Ing. LN</td>
<td>A A Q Y L A</td>
</tr>
</tbody>
</table>
Macaques inoculated with SHIV\textsubscript{VifAAQYLA} developed antibody responses against SHIV

At necropsy, we analyzed the plasma for the presence of immunoprecipitating antibodies against SHIV\textsubscript{KU-1bMC33}. All three macaques developed immunoprecipitating antibody responses against SHIV\textsubscript{KU-1bMC33}, although macaque RCS10 developed significantly lower antibody response compared to the other two macaques (Figure 19). In contrast, a macaque inoculated with parental SHIV\textsubscript{KU-1bMC33} (RRH10) did not develop antibodies to the virus, which is common for macaques that develop severe CD4\textsuperscript{+} T cell loss during the acute phase (<4 weeks) following inoculation with pathogenic X4 SHIV (Figure 19).

**Histological examination of tissues**

Tissues from macaques were examined for the presence of histological lesions consistent with pathogenic X4 SHIV infection (Stephens et al., 2002). Histological examination of tissues from macaque RRH10, which was inoculated with SHIV\textsubscript{KU-1bMC33}, revealed severe lymphoid depletion in both the thymus and lymph nodes and mild lymphoid depletion in the spleen (Figure 20). The lymphoid depletion observed in the macaques is consistent of that observed following inoculation with the parental SHIV\textsubscript{KU-1bMC33} (Stephens et al., 2002). The macaques inoculated with SHIV\textsubscript{VifAAQYLA} did not exhibit lesions in any of the 13 visceral organs and CNS. Micrographs of sections from the thymus, mesenteric lymph node, and spleen from macaque RAK10 are shown along with micrographs of histological sections from an uninfected macaque in Figure 20.
Figure 19. Macaques inoculated with SHIV_{VifAAQYLA} developed antibody responses against SHIV. C8166 cells were inoculated with SHIV_{KU-1bMC33} for 5 days, starved in methionine/cysteine-free media and radiolabeled overnight with $^{35}$S-methionine/cysteine. The culture medium was harvested and used in immunoprecipitation reactions with plasma from RAK10, RCS10, RPL10 and RRH10 as described in the Experimental Methods. The immunoprecipitates were washed with 1x RIPA buffer, boiled in 2x sample reducing buffer and proteins separated on 10% SDS-PAGE gels. (Lane 1) SHIV proteins immunoprecipitated using supernatant from an uninfected culture. (Lane 2) SHIV proteins immunoprecipitated from a positive control plasma sample. (Lane 3) SHIV proteins immunoprecipitated using RAK10 plasma. (Lane 4) SHIV proteins immunoprecipitated using RCS10 plasma. (Lane 5) SHIV proteins immunoprecipitated using RPL10 plasma. (Lane 6) SHIV proteins immunoprecipitated using RRH10 plasma. The molecular weight standards are displayed on the left and specific SHIV proteins are shown on the right.
Figure 20. Histopathology associated with SHIV_{VifAAQYLA} infection of macaques. Hematoxylin and eosin stains of section from the thymus (A-C), mesenteric lymph node (Panels D-F) and spleen (Panels G-I) from a SHIV_{VifAAQYLA} inoculated macaque RAK10 (Panels A, D, and G), an uninfected macaque (Panels B, E and H) and a SHIV_{KU-1bMC33} infected macaque RRH10 (Panels C, F, and I).
Macaques inoculated with SHIV_{VifAAQYLA} had decreased tissue distribution of replicating virus

Since the plasma viral loads were decreased in comparison to macaques inoculated with parental SHIV_{KU-1bMC33}, we were interested in determining if this effect was due to the infection of select cell populations in one or more organs, which may not express rhesus A3G. First, we analyzed DNA isolated from 13 visceral organs from SHIV_{VifAAQYLA} macaques for the presence of viral gag sequences by nested DNA PCR. We found that 12 of the 13 visceral organs from a SHIV_{KU-1bMC33} inoculated macaque (RRH10) were positive for viral gag sequences in DNA samples (Figure 21). For macaques inoculated with SHIV_{VifAAQYLA}, we found that 6 of the 13, 8 of the 13, and 12 of the 13 organs from macaques RAK10, RCS10 and RPL10, respectively, were positive for viral gag sequences by nested DNA PCR (Figure 21). The majority of the organs that were gag positive were lymphoid organs. We next assessed RNA isolated from these tissues for the presence of viral gag RNA sequences using RT-PCR. The results of this analysis are shown in Figure 22. We found that 5 of 13, 6 of 13, and 9 of 13 visceral organs from macaques RAK10, RCS10 and RPL10 were positive for viral gag RNA sequences, respectively. A macaque inoculated with SHIV_{KU-1bMC33} (RRH10) had 11 of the 13 organs positive for the presence of gag RNA sequences. Taken together, these results indicate that the SHIV_{VifAAQYLA} infection was less widespread in 2 or the 3 macaques compared to a macaque inoculated with parental SHIV_{KU-1bMC33}. 
Figure 21. Macaques inoculated with SHIV_{VifAAQYLA} had a decreased tissue distribution of proviral DNA. DNA was isolated from different organs as indicated at necropsy and amplified using nested DNA PCR and oligonucleotides specific for gag. Samples were run on a 1.5% agarose gel, stained with ethidium bromide, and photographed. Shown are the results of the nested DNA PCR reaction using DNA from macaques RRH10 (Panel A), RAK10 (Panel B), RCS10 (Panel C) and RPL10 (Panel D). The tissues DNAs analyzed are noted above each lane and the 100 base pair marker is indicated on the right. The positive control was gag amplified from lymph node tissue obtained from macaque 50O that died of neuroAIDS (McCormick-Davis et al., 2000) and the negative control was spleen tissue DNA from an uninfected macaque.
Figure 22. Macaques inoculated with SHIV\textsubscript{VifAAQYLA} had a decreased tissue distribution of viral RNA. RNA was isolated from 13 different visceral organs obtained at necropsy. The RNA was DNase I treated to remove residual DNA and used in a nested RT-PCR reaction with oligonucleotides specific for \textit{gag}. Samples were run on a 1.5% agarose gel, stained with ethidium bromide and photographed. The results are shown as follows: Macaques RRH10 (Panel A), RAK10 (Panel B), RCS10 (Panel C) and RPL10 (Panel D). The tissue RNAs analyzed are noted above each lane. The positive control was amplified from lymph node tissue obtained from macaque 50O that died of neuroAIDS (McCormick-Davis et al., 2000). The negative control was spleen RNA isolated from an uninfected macaque. The 100 bp marker is indicated on the right.
Sequence changes in *vif, vpu, env* and *nef* in primary and secondary lymphoid organs at necropsy

Previous research has shown that HIV-1Δ*vif* incorporates select A3 proteins into viral particles leading to cytosine deamination in the minus strand during DNA synthesis (Sheehy et al., 2002; Bishop et al., 2004). Hypermutation in HIV-1 reverse transcripts display a high frequency of G-to-A substitutions in the plus strand of DNA resulting in inhibition of virus replication (Harris et al., 2003; Mangeat et al., 2003; Zhang et al., 2003). We examined the extent of G-to-A mutations in viral DNA amplified from one primary lymphoid organ (thymus) and several secondary lymphoid organs (mesenteric, axillary, and inguinal lymph nodes) from three macaques inoculated with SHIV*VifAAQYLA* and one macaque inoculated with SHIV*KU-1bMC33* (RRH10). We analyzed sequences from the *vif, vpu, env,* and *nef* genes. The results of the sequence analysis are shown in Table III. These analyses showed that: a) with the exception of *env,* the thymus consistently had the lowest percentage of G-to-A substitutions; b) not all lymphoid organs had significant G-to-A substitutions; and c) greater than 85% of the G-to-A mutations were in the context of TC and not CC. The sequence of *vpu* amplified from the lymphoid tissues of macaque RPL10 is shown as an example in Figure 23. Macaque RRH10 had the fewest G-to-A substitutions in all of the lymphoid organs, which is similar to other macaques inoculated with SHIV*KU-1bMC33* (unpublished observations). The percentage of mutations was approximately <0.2-0.41% of the nucleotides analyzed for macaque RRH10. The highest percentage of G-to-A substitutions was found in the *nef* gene of macaque RCS10 inoculated with SHIV*VifAAQYLA* (5.0%).
Table III. Results of viral DNA sequence analysis from visceral tissues at necropsy
<table>
<thead>
<tr>
<th>Macaque/ Gene</th>
<th>Tissue</th>
<th>RRH10</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Thymus</td>
<td>1/466 (0.21%)</td>
<td></td>
<td>1/466 (0.21%)</td>
<td></td>
<td>1/466 (0.21%)</td>
<td></td>
</tr>
<tr>
<td>vif</td>
<td>MLN</td>
<td>1/466 (0.21%)</td>
<td></td>
<td>1/466 (0.21%)</td>
<td></td>
<td>1/466 (0.21%)</td>
<td></td>
</tr>
<tr>
<td>vpu</td>
<td>ALN</td>
<td>2/246 (0.0%)</td>
<td></td>
<td>1/246 (0.4%)</td>
<td></td>
<td>1/246 (0.4%)</td>
<td></td>
</tr>
<tr>
<td>env</td>
<td>ILN</td>
<td>2/490 (0.41%)</td>
<td></td>
<td>0/490 (0.0%)</td>
<td></td>
<td>0/480 (0.0%)</td>
<td></td>
</tr>
<tr>
<td>nef</td>
<td></td>
<td>1/400 (0.25%)</td>
<td></td>
<td>1/400 (0.25%)</td>
<td></td>
<td>2/400 (0.50%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Macaque/ Gene</th>
<th>Tissue</th>
<th>RAK10</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Thymus</td>
<td>0/466 (0.00%)</td>
<td></td>
<td>16/466 (3.4%)</td>
<td></td>
<td>7/400 (1.50%)</td>
<td></td>
</tr>
<tr>
<td>vif</td>
<td>MLN</td>
<td>1/246 (0.4%)</td>
<td></td>
<td>4/246 (0.4%)</td>
<td></td>
<td>4/246 (1.63%)</td>
<td></td>
</tr>
<tr>
<td>vpu</td>
<td>ALN</td>
<td>12/490 (2.44%)</td>
<td></td>
<td>10/490 (2.04%)</td>
<td></td>
<td>12/490 (2.44%)</td>
<td></td>
</tr>
<tr>
<td>env</td>
<td>ILN</td>
<td>2/400 (0.50%)</td>
<td></td>
<td>1/400 (0.25%)</td>
<td></td>
<td>17/400 (4.25%)</td>
<td></td>
</tr>
<tr>
<td>nef</td>
<td></td>
<td>0/466 (0.0%)</td>
<td></td>
<td>1/466 (0.21%)</td>
<td></td>
<td>8/466 (3.25%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Macaque/ Gene</th>
<th>Tissue</th>
<th>RCS10</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Thymus</td>
<td>0/466 (0.0%)</td>
<td></td>
<td>1/466 (0.21%)</td>
<td></td>
<td>8/466 (3.25%)</td>
<td></td>
</tr>
<tr>
<td>vif</td>
<td>MLN</td>
<td>1/246 (0.4%)</td>
<td></td>
<td>1/246 (0.4%)</td>
<td></td>
<td>4/246 (1.62%)</td>
<td></td>
</tr>
<tr>
<td>vpu</td>
<td>ALN</td>
<td>10/490 (2.04%)</td>
<td></td>
<td>8/490 (1.63%)</td>
<td></td>
<td>10/490 (2.04%)</td>
<td></td>
</tr>
<tr>
<td>env</td>
<td>ILN</td>
<td>1/400 (0.25%)</td>
<td></td>
<td>20/400 (5.00%)</td>
<td></td>
<td>2/400 (0.5%)</td>
<td></td>
</tr>
<tr>
<td>nef</td>
<td></td>
<td>0/466 (0.0%)</td>
<td></td>
<td>1/466 (0.21%)</td>
<td></td>
<td>8/466 (0.21%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Macaque/ Gene</th>
<th>Tissue</th>
<th>RPL10</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Thymus</td>
<td>3/466 (0.64%)</td>
<td></td>
<td>6/466 (1.28%)</td>
<td></td>
<td>9/466 (1.93%)</td>
<td></td>
</tr>
<tr>
<td>vif</td>
<td>MLN</td>
<td>3/246 (1.2%)</td>
<td></td>
<td>6/246 (2.4%)</td>
<td></td>
<td>9/246 (3.70%)</td>
<td></td>
</tr>
<tr>
<td>vpu</td>
<td>ALN</td>
<td>9/490 (1.83%)</td>
<td></td>
<td>8/490 (1.63%)</td>
<td></td>
<td>8/490 (1.63%)</td>
<td></td>
</tr>
<tr>
<td>env</td>
<td>ILN</td>
<td>1/400 (0.25%)</td>
<td></td>
<td>2/400 (0.5%)</td>
<td></td>
<td>6/400 (1.25%)</td>
<td></td>
</tr>
<tr>
<td>nef</td>
<td></td>
<td>3/466 (1.72%)</td>
<td></td>
<td>8/466 (3.3%)</td>
<td></td>
<td>8/466 (3.3%)</td>
<td></td>
</tr>
</tbody>
</table>

The ratio of the number of G-to-A nucleotide substitutions/number of bases sequenced is shown.

a The percentage of G-to-A substitutions are in parentheses.
Figure 23. Sequence of the \textit{vpu} gene amplified from lymphoid tissues of macaque RPL10. Shown on the top line is the input sequence of \textit{vpu} from SHIV\textsubscript{KU-1bMC33}. The (.) represents identity with the nucleotide sequence shown above. The underlined sequences represent the context in which the G-to-A substitutions occurred.
Pooled plasma from two macaques inoculated with SHIV\textsubscript{VifAAQYLA} contained low levels of infectious virus when transfused into a naïve macaque

In order to determine if infectious virus was present in plasma that was not detected in our infectivity assays, plasma from macaques RAK10 and RPL10 was pooled and intravenously inoculated into a naïve macaque, I95. We were able to detect viral \textit{gag} sequences in DNA isolated from PBMC at 0, 1, 2, 3, 4, 5, 6, and 20 weeks post-inoculation (Figure 25). We also assessed plasmas from macaque I95 for viral loads using real-time PCR from weeks 0 to 20. We were unable to detect viral burdens in macaque I95. The level of circulating CD4\textsuperscript{+} T cells did not decrease from the pre-inoculation levels (Figure 24). We were unable to detect antibodies in the plasma from this macaque at week 9 (data not shown). At 20 weeks post-inoculation, macaque I95 was sacrificed. DNA was isolated from PBMC at various time points throughout the duration of infections and from 10 visceral organs and used in a nested PCR to detect viral \textit{gag} sequences. Viral \textit{gag} sequences were amplified from all 10 visceral organs examined (Figure 26). We also isolated RNA from the same organs for use in RT-PCR to detect viral \textit{gag} RNA sequences. The viral RNA was detected in six organs (liver, kidney, axillary lymph node, small intestine (ileum), spleen, and tonsil) (Figure 27). This indicates that some level of viral replication was ongoing although we were unable to isolate infectious virus on CEM-SS and SupT1 cells (data not shown). Taken together, these results indicate that pooled plasma contained only low levels of infectious virus that were controlled by the immune system of macaque I95.
Figure 24. Viral loads and circulating CD4$^+$ T cell levels in macaque I95 that was intravenously inoculated with pooled plasma from macaques RAK10 and RPL10 obtained at necropsy. Blood was obtained at 0 (just prior to inoculation), 1, 2, 3, 4, 5, 6 and 20 weeks after inoculation. Results of the real-time PCR for viral gag sequences from plasma samples (▲; scale to left) and circulating CD4$^+$ T cell levels (●; scale to right).
Figure 25. Presence of viral DNA sequences in PBMC from macaque 195. PBMC were purified over Ficoll-Hypaque gradients and total cellular DNA was isolated. DNA was used in a nested DNA PCR for viral gag. (Lane 1) 100 base pair marker. (Lane 2) Positive control, week 1 PBMC DNA from macaque RRH10 (inoculated with SHIV\textsubscript{KU-1\textsubscript{bMC33}}. (Lane 3-10) Represent gag sequences amplified from PBMC isolated from I95 at 0, 1, 2, 3, 4, 5, 6, and 20 weeks post-inoculation, respectively.
Figure 26. Distribution of viral DNA gag sequences in tissues from macaque I95. At necropsy tissues were harvested, DNA isolated and a nested PCR for gag run. (Lane 1) Positive control, lymph node tissue DNA from macaque 500 that died of neuroAIDS (McCormick-Davis et al., 2000). (Lane 2) Negative control, spleen tissue DNA from an uninfected macaque. (Lane 3) Liver. (Lane 4) Lung. (Lane 5) Kidney. (Lane 6) Mesenteric lymph node. (Lane 7) Axillary lymph node. (Lane 8) Inguinal lymph node. (Lane 9) Small intestine (ileum). (Lane 10) Spleen. (Lane 11) Thymus. (Lane 12) Tonsil. (Lane 13) 100 base pair marker.
Figure 27. Distribution of viral RNA gag sequences in tissues from macaque I95. At necropsy, tissues were harvested for total RNA and a nested RT-PCR for gag was run. (Lane 1) 100 base pair marker. (Lane 2) Positive control, lymph node tissue RNA from macaque 50O that died of neuroAIDS (McCormick-Davis et al., 2000). (Lane 3) Negative control, spleen tissue DNA from an uninfected macaque. (Lane 4) Liver. (Lane 5) Lung. (Lane 6) Kidney. (Lane 7) Mesenteric lymph node. (Lane 8) Axillary lymph node. (Lane 9) Inguinal lymph node. (Lane 10) Small intestine (ileum). (Lane 11) Spleen. (Lane 12) Thymus. (Lane 13) Tonsil.
Discussion

The results presented in this publication are the first to evaluate the role of domain-specific mutations in the BC box domain of the \textit{vif} gene using a non-human primate model of HIV-1/AIDS. Previous studies have shown that the BC box domain of HIV-1 Vif (SLQYLALAAL) interacts with elongin C, which contains the consensus sequence $(A, P, S, T)_1 L_2 x x x (C, A, S)_6 x x x (A, I, L, V)_10$ with the leucine at position 2 being invariant (Luo et al., 2005; Mehle et al., 2004). The invariant leucine at position 2 of the BC box domain is critical because mutation of this amino acid to an alanine prevents the interaction between Vif and elongin C preventing assembly of the E3 ubiquitin Ligase complex (Kobayashi et al., 2005; Yu et al., 2003; Yu et al., 2004a, 2004b, 2004c). Since the S147 and L148 amino acid residues of the SLQYLALAAL domain are highly conserved among primate lentivirus Vif proteins, we chose to mutate these amino acid residues to alanine residues. The results from this study indicate that macaques inoculated with SHIV\textit{VifAAQYLA} developed no significant loss of circulating CD4\(^+\) T cells and had a 100-fold lower plasma viral loads during the acute phase of infection, when compared to macaques inoculated with SHIV\textit{KU-1bMC33}. Additionally, our analyses revealed fewer tissues positive for viral \textit{gag} DNA and RNA sequences in 2 of the 3 macaques inoculated with SHIV\textit{VifAAQYLA}, suggesting that SHIV\textit{VifAAQYLA} was controlled by the host. This could be due to the presence of susceptible cells expressing one or more of the A3 proteins that could restrict a virus containing mutations such as these. The presence of plasma viral loads in these macaques indicates that virus may have persisted in populations of cells that either did not express A3 proteins or expressed these proteins at very low levels.
Thus, a major question that emerged was, “Are the viruses produced from these cellular reservoirs infectious, replication competent viruses?” In order to detect small amounts of virus that we were unable to see in our infectious centers assay (data not shown), we inoculated a naïve macaque, I95, with plasma pooled from 2 of the 3 macaques inoculated with SHIV\textsubscript{VifAAQYLA}. Macaque I95 revealed the presence of virus in PBMCs between weeks 1 to 20 and in 10 tissues (liver, lung, kidney, mesenteric lymph node, axillary lymph node, inguinal lymph node, ileum, spleen, thymus, and tonsil) examined at necropsy, although the plasma viral loads were below the limits of detection (Figure 25). These results suggested that plasma from the SHIV\textsubscript{VifAAQYLA} inoculated macaques had infectious virus, but the macaques readily controlled the virus.

While none of the three macaques inoculated with SHIV\textsubscript{VifAAQYLA} exhibited histological lesions, there were two incongruent observations. First, macaque RPL10 had the lowest viral loads but also had the most widespread visceral tissue distribution of viral gag sequences. While the reason for this result is unclear, it could be explained by a greater viral distribution during the early stage of infection. Second, macaques RCS10 developed less immunoprecipitating antibodies compared with macaques RAK10 and RPL10. Although the answer to this question is currently unknown, if could relate to the inability of RCS10 to select for the A147T amino acid substitution. Perhaps the A147T substitution permitted more replication in macaques RAK10 and RPL10 providing a better antigenic stimulus. It should also be noted that SHIV\textsubscript{VifTAQYLA} from RAK10 and RPL10 could not be isolated. Therefore, we do not know if a virus with the \textsuperscript{146}TAQYLALAA\textsuperscript{155} motif is capable of causing CD4+ T cell loss, histological lesions or high viral burdens.
Our sequence analysis indicated that not all viral genes amplified from the lymphoid tissues contained G-to-A substitutions above background levels. For the three macaques inoculated with SHIV\textsubscript{VifAAQYLA}, a total of four lymph tissues (thymus, mesenteric, axillary, and inguinal lymph nodes) and four viral genes (\textit{vif}, \textit{vpu}, \textit{env}, and \textit{nef}) were assessed (Table III). Approximately half (25) of the 48 sequences contained G-to-A substitutions that were above 1.2% with the majority of the substitutions being in the context of 5'-TC. Potential reasons for this may include that the aliquot of tissue examined may have contained archived virus from the initial round of infection, the viruses amplified were from a cell type that did not express A3G, or that the domain-specific mutations may have crippled but not completely abolished the antiviral activity of Vif. Alternatively, it is possible that both deaminase-dependent and deaminase-independent mechanisms of restriction may be present to decrease the replication of SHIV\textsubscript{VifAAQYLA} in macaques. The context of the G-to-A substitutions identified in viral sequences can provide insight into which A3 members are involved in cytidine deamination. For instance, human A3G is known to cause cytidine deamination with a base preference of 5'-CC, while A3F, A3H and A3A have a preference for 5'-TC (Beale et al., 2004; Langolis et al., 2005; Liddament et al., 2004; Thiel et al., 2010). However, A3D has a preference for 5'-T/A-T/A-C-G/T and A3B and A3C favor both 5'-TC and 5'-CC (Dang et al., 2007; Doehle et al., 2005; Yu et al., 2004). This indicates that most likely an A3 protein other than A3G is causing the majority of the mutations we observed.

Attenuated viruses have been used to prevent many viral diseases in humans (smallpox, poliovirus, measles, mumps, rubella, chickenpox etc.). It is known that live,
attenuated virus vaccines generally result in better immune responses (both humoral and cell mediated) that are longer lived when compared to killed virus vaccines. However, the major drawback of attenuated virus vaccines is that live viruses are genetic elements and replication can introduce mutations that result in reversion of the attenuated phenotype to a pathogenic phenotype. The prime example of this is the live attenuated poliovirus vaccine (Kew et al., 2005). The major concern for the use of attenuated lentiviruses has been that it involved the removal of “non-essential” viral genes such as vpr, vpx, nef or vpu, which could allow for low level virus replication and the accumulation of compensatory mutations in other gene and potentially a reversion to a pathogenic phenotype. The work by Daniel and colleagues (1992) showed that deletion of the nef from the pathogenic SIV<sub>mac</sub>239 resulted in an attenuated virus that resisted challenge against pathogenic parental virus. However, the hope that an attenuated vaccine approach was soon diminished by the studies showing that inoculation of a SIV<sub>mac</sub>239 derivative with deletions in vpr, nef and the negative regulatory element (NRE) could cause AIDS in neonatal macaques (Baba et al., 1995). This was followed with a study showing that the same virus could cause disease in both neonatal and adult macaques (Baba et al., 1999). Despite these setbacks, the vaccines against SIV (or HIV-1) infection that have shown the most promise have been those that have deleted or inactivated one or more genes of the virus (Koff et al., 2006). In contrast to Vpr, Vpx, Vpu and Nef, the Vif protein is required for replication in CD4<sup>+</sup> T cells and macrophages, the two major cell populations that HIV-1 infects productively (Chiu and Greene, 2008). This suggests that mutations that target one or more conserved (and functional) domains within Vif may permit limited replication and
generation of effective immune responses against the virus prior to “self inactivation” through the activities of various APOBEC3 family members and may represent a novel means to attenuate lentiviruses for candidate vaccines.

**Experimental Methods**

**Cells, Plasmids, and Viruses**

The CEM and CEM-SS lymphocyte cell lines obtained from the NIH AIDS Research and Reagents Program were used for the production of viral stocks, the measure of infectivity and the cytopathicity of the viruses used in this study (Salter et al., 1985; Foley et al., 1965; Nara et al., 1987; Nara et al., 1988). CEM, CEM-SS, and C8166 cell lines were maintained in RPMI-1640 supplemented with 10 mM Hepes buffer pH 7.3, 10% fetal bovine serum (R10FBS), 5 µg/mL gentamicin, and 1% penicillin/streptomycin. The derivation of SHIV\textsubscript{KU-1bMC33} has been previously described (McCormick-Davis et al., 2000; Stephens et al., 2002).

**Proviral DNA Plasmid Construction**

For the construction of SHIV\textsubscript{VifAAQYLA}, a PacI/SphI fragment (nucleotides 4609 to 5901) from the p5'SHIV-4 was subcloned into the pGEM3Zf (+) vector. The serine at amino acid position 147 of Vif was substituted to an alanine using oligonucleotides (only sense strand shown) 5’-AAGTACCAGGTACCCCTACAGTACTTA-3’ and the Quick-Change Mutagenesis Kit (Stratagene) according to the manufacturer’s instructions. The resulting plasmid, pGEM3Zf (+)-Paci/SphI Vif\textsubscript{AAQYLA}, was used to
change the leucine at amino acid position 148 of Vif to an alanine using oligonucleotides (sense strand shown) 5’-ACCAGGTACCAGCCGCACAGTACTTAGCAC-3’ and the Quick-Change Mutagenesis Kit (Stratagene) according to the manufacturer’s instructions. The resulting plasmid, pGEM3Zf (+)-VifAAQYLA, was purified and digested with Pacl/SphI. The isolated fragment was purified and subcloned into p5’SHIV-4 to generate p5’SHIV\textsubscript{VifAAQYLA}. The proviral plasmid DNA was sequenced to determine if the desired mutations were introduced as expected. To generate virus, p5’SHIV\textsubscript{VifAAQYLA} and p3’SHIV were digested with SphI and the two plasmids were ligated together with T4 DNA Ligase (Promega). The resulting ligation was transfected into CEM-SS cells as previously described (Stephens et al., 2002; Hout et al., 2005; 2006). Stocks of SHIV\textsubscript{KU-1bMC33} and SHIV\textsubscript{VifAAQYLA} were harvested, titrated onto CEM-SS cells to determine the TCID\textsubscript{50}/mL and stored at -86°C.

**Replication of SHIV\textsubscript{VifAAQYLA} in both non-permissive and permissive cell lines**

The replication of SHIV\textsubscript{VifAAQYLA} was assessed in CEM (A3G/F positive) and CEM-SS (A3G/F negative) cell lines. One million cells were infected with equivalent levels (50 ng) of parental SHIV\textsubscript{KU-1bMC33} or SHIV\textsubscript{VifAAQYLA} for 4 h at 37°C. The cells were centrifuged and washed 3 times to remove the inoculum and incubated in fresh medium at 37°C for 9 days. Aliquots of culture supernatants were obtained at 0, 1, 3, 5, 7 and 9 days post-inoculation. The level of p27 was assessed using a commercially available p27 antigen capture assay kits (Zeptometrix).

**Virion Incorporation Assay**
Plasmids containing the genomes of viruses SHIV$_{KU-1bMC33}$ and SHIV$_{VifAAQYLA}$ were co-transfected with human HA-APOBEC3G (HA-hA3G) into 293 cells using polyethylenimine transfection reagent (PEI; ExGen-500, Fermentas) along with a plasmid expressing HA-hA3G. At 48 hours, virus containing supernatants were harvested and clarified by low speed centrifugation. The clarified supernatant was then subjected to ultracentrifugation to pellet the virus (SW41 rotor 247,000 xg, 1 h). The pellet was resuspended in 1x PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$·2H$_2$O, 1.76 mM KH$_2$PO$_4$; pH 7.4). An aliquot was saved to determine the p27 content by a commercially available antigen capture assay (Zeptometrix). The remaining sample was boiled in 2x sample reducing buffer (125 mM Tris pH 6.8, 4% SDS, 10% glycerol, 0.006% bromophenol blue, 1% β-mercaptoethanol) and boiled for 5 minutes. Equivalent amounts of p27 were loaded on a 12% SDS-PAGE gel and transferred to PVDF membrane. HA-hA3G proteins were detected by Western blotting using an antibody directed against the HA tag (HA-probe, Santa Cruz) of HA3G. The blot was stripped in 1x stripping buffer (62.5 mM Tris-HCl, pH6.8 and 2% SDS in 1x PBS, pH 7.4) and reprobed using a rabbit polyclonal antibody specific for p27.

Macaques analyzed in this study

Five 1-1.5 year old pig-tailed macaques (Macaca nemestrina: 2000; CM4G; CM4K; RRH10) were intravenously inoculated with 1 ml of undiluted SHIV$_{KU-1bMC33}$ containing 10$^4$ TCID$_{50}$ per ml (titered in CEM cells). Three additional pig-tailed macaques (RPL10; RCS10; RAK10) were inoculated with 10$^4$ TCID$_{50}$ SHIV$_{VifAAQYLA}$ (titered in CEM-SS cells). Twenty-four weeks post-inoculation, 2 ml of plasma was
pooled from macaques RPL10 and RAK10 and used to inoculate a naïve macaque, I95. The animals were all housed in the AAALAC-approved animal facility at the University of Kansas Medical Center. Blood was collected weekly for the first 4 weeks, then at 2-week intervals for the next month, and thereafter at monthly intervals.

Circulating CD4+ T cells

The levels of circulating CD4+ T cells was assessed at 0, 1, 2, 3, 4, 6, 8, 12, 16, 20, 24, 28 weeks post-inoculation using FACs (Bectin Dickinson) analysis. T cell subsets were labeled with OKT4 (CD4; Ortho Diagnostics Systems, Inc.), SP34 (CD3; Pharmigen), or FN18 (CD3; Biosource International) monoclonal antibodies. The number of circulating CD4+ T cells were determined per µl. As a control for FACS analysis, T cell subsets from an uninfected macaque were always performed.

Processing of tissue samples at necropsy

At the time of euthanasia (28 weeks for those macaques inoculated with SHIV\text{V\text{II}AAQYLA} and 20 weeks for macaque I95), all animals in this study were anesthetized by administration of 10-mg/kg ketamine (IM) followed by intravenous administration of sodium Phenobarbital at 20-30 mg/kg. A laparotomy was performed and the animal exsanguinated by aortic canulation and perfused with one liter of cold Ringer's saline. All aspects of the animal studies were performed according to the institutional guidelines for animal care and use at the University of Kansas Medical Center. Lymphoid and non-lymphoid tissues (brain, heart, lungs, liver, kidney, pancreas, salivary glands, liver, spleen, thymus, small intestine, colon, tonsils, and
mesenteric, inguinal and axillary lymph nodes) were harvested in HBSS (HBSS; 0.137 M NaCl, 5.4 mM KCl, 0.25 mM Na$_2$HPO$_4$, 0.44 mM KH$_2$PO$_4$, 1.3 mM CaCl$_2$, 1.0 mM MgSO$_4$, 4.2 mM NaHCO$_3$) and aliquots of tissues were either fixed for histological analysis or snap frozen for DNA and RNA assays. Cells were isolated from the thymus, spleen, and mesenteric lymph node and used in an infectious centers assay to determine if infectious virus is present at the time of necropsy.

**Histology**

Tissues obtained from necropsy listed above were fixed in 10% neutral buffered-formalin, subdivided into smaller blocks, and processed for paraffin embedding. Five µm sections were stained with eosin and hematoxylin. The histological samples were examined for morphological abnormalities and lesions consistent with pathogenic X4 SHIV infection (Stephens et al., 2002).

**Plasma viral loads**

Plasma viral RNA loads were determined on RNA extracted from 0.5 ml of EDTA-treated plasma. Virus was pelleted using ultracentrifugation (Beckman SW55Ti, 110,000 x g, 2 h) and RNA extracted using the Qiagen viral RNA kit (Qiagen). RNA samples were analyzed by real-time RT-PCR using gag primers and a 5'FAM and 3'TAMRA labeled Taqman probe that was homologous to the SIV gag gene as previously described (Hofmann-Lehmann et al., 2000). Standard curves were prepared using a series of 12 ten-fold dilutions of viral RNA of known concentration. The sensitivity of the assay was 100 RNA equivalents per ml. Samples were analyzed in
triplicate and the number or RNA equivalents were calculated per ml of plasma. The limit of detection was 100 copies.

**Analysis of tissues for viral DNA and RNA**

DNA was extracted from the visceral organs listed above as previously described (Hill et al., 2008) and used in nested PCR to amplify SIV *gag* sequence. The 240 base pair amplified product was visualized through a 1.5% agarose gel stained with ethidium bromide. Viral RNA, which is indicative of actively replicating virus, was extracted from approximately 30 mg of each visceral or CNS tissue using the RNAEasy kit (Qiagen) and the manufacturer’s instructions. RNA samples were digested with DNase I for 30 minutes and run on agarose formaldehyde gels to check for the presence of contaminating DNA. RT-PCR amplification of the extracted RNA was performed using the Titan One RT-PCR kit (Roche) and the manufacturer’s instructions. Each reaction used one microgram of total RNA and was amplified using oligonucleotide primers specific for the SIV<sub>mac</sub>239 *gag* gene. Each RT PCR reaction contained 50 ng of RNA, 0.1 mM dNTP mix, 5 mM DTT solution, 5U RNase inhibitor, 0.4 µM of each primer, 5X RT-PCR buffer, and 1U of Enzyme mix. The oligonucleotides used for the first round of amplification were 5’-CGTCATCTGGTGCTATCACG-3’ (sense) and 5’-CTGATTAATGTCATAGGGGGTGC-3’ (antisense), which are complementary to bases 1343-1362 and 1636-1658, respectively. The nested primers used were 5’-CACGCAGAAGAGAAAGTGAAACAC-3’ (sense) and 5’-GGTGCAACCTTCTGACAGTGC-3’ (antisense), which are correspond to bases 1359-1382 and 1620-40, respectively. The reactions were performed with an Applied
Systems 2720 Thermal Cycler using the following thermal profile: 42°C for 30 min, 1 cycle; 94°C for 2 min, 1 cycle; 94°C for 30 s, 55°C for 30 s, 68°C for 45 s, 10 cycles; 94°C for 30 s, 55°C for 30 s, and 68°C for 2 min; 25 cycles. One microliter of the initial reaction mixture was then added to a nested PCR mixture containing gag primers and performed with the following thermal profile: 95°C for 10 min, 1 cycle; 30 s, 55°C for 30 s, and 68°C for 2 min; 35 cycles. The amplified gag fragment is 281 base pairs. β-actin was amplified from all of the tissue RNA samples as a control to verify the RNA integrity.

**DNA sequence analysis of *vif*, *vpu*, *nef*, and *env***

In order to assess the sequence of viral genes, PBMC and tissue DNA was extracted using the QIAamp DNA mini kit (Qiagen) according to the manufacturer’s instructions. One μg of isolated genomic DNA was used in a nested DNA polymerase chain reaction using rTaq (Takara) following the manufacturer’s instructions and specific nucleotides listed in Table IV. Each PCR reaction contained 1x PCR buffer (Mg²⁺-free), 4 mM MgCl₂, 0.25 mM of each deoxynucleotide triphosphate, 500 nM of each primer, and 1U of rTaq DNA polymerase. The PCR reactions were performed using an Applied Systems 2720 Thermal Cycler with the following thermal profile: 95°C for 2 min; 1 cycle; 95°C for 30 s, 48°C for 30 s, 65°C for 2 min, 35 cycles; 65°C for 7 min. The PCR products were pooled from three separate reactions and separated by electrophoresis on a 1.5% agarose gel, isolated, and directly sequenced. Cycled sequencing reactions were constructed using the BigDye Terminator Cycle Sequencing Ready Reaction kit with AmpliTaq DNA Polymerase, FS (PE Applied Biosystems). Sequence detection
Table IV. Oligonucleotides used in rTaq PCR to amplify SHIV genes.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Fragment Size (base pairs)</th>
<th>PCR Reaction</th>
<th>Forward (5' to 3')</th>
<th>Reverse (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>nef</td>
<td>792</td>
<td>Round 1</td>
<td>CCTAGAAGAATAAGACAGGCTTGG</td>
<td>GGCTGACAAGAAGGAAACTGCG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Round 2</td>
<td>GGAGCTATTTCCATGAGCGGTCC</td>
<td>GGCCCTTTCCGGTTAGCCCTTC</td>
</tr>
<tr>
<td>vpu</td>
<td>249</td>
<td>Round 1</td>
<td>CCTAGACTAGAGCCCTGGAAGCATTCC</td>
<td>GTACCTCTCTGATCATATGCTTATGAGCAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Round 2</td>
<td>TTAGGCATCTCCTATGGGCGAGGAAGAG</td>
<td>CACAAAATAGAGTGGGTGGTTCTTCCT</td>
</tr>
<tr>
<td>vif</td>
<td>645</td>
<td>Round 1</td>
<td>GGCTAAAATTATCAAAGATTATGGGAGG</td>
<td>GGTCGACATCCCTTTTCTCATACAGC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Round 2</td>
<td>GGAGGGAGGAAAGAGGTGATAGCAG</td>
<td>GGCTACGACCTCCCTTTGAGCC</td>
</tr>
<tr>
<td>env</td>
<td>810</td>
<td>Round 1</td>
<td>GGTAAAAATGATACTAATAAGAAAGAG</td>
<td>CAGACCTCTCTGTAGACTGCC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Round 2</td>
<td>GCAGAAGACAGTGCAATGAGAG</td>
<td>CAGCACTTGAGTTGTAATCTACAG</td>
</tr>
<tr>
<td></td>
<td>765</td>
<td>Round 1</td>
<td>GCTGTTTTGGCATTCTAAAGTG</td>
<td>GCTTCCTCGCTCTCCCAAAGACCC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Round 2</td>
<td>GTAATAAATGAAGCCTTCAATGG</td>
<td>GGAACAAAGCTTCCTATTC</td>
</tr>
<tr>
<td>gag</td>
<td>240</td>
<td>Round 1</td>
<td>CGTGCAACCTTTCGATCTCACG</td>
<td>CTGATTATGTATGAGGGGTGC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Round 2</td>
<td>TCATGTTTGAGACCTTCAACACCACAG</td>
<td>CCAAGAAGAAGGCTGGGAGAGTAGG</td>
</tr>
<tr>
<td>β-actin</td>
<td>300</td>
<td>Round 1</td>
<td>TCATGTTTGGAGACCTTCAACACCACAG</td>
<td>CCAAGGAAGAAGGCTGGGAGAGTAGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Round 2</td>
<td>CCCAGCCATGTAGTGGCTATCC</td>
<td>GCCTACGGGCAAGGAAACCAGCTCA</td>
</tr>
</tbody>
</table>
was conducted with an Applied Biosystems 377 Prism XL automated DNA sequencer and visualized using the ABI Editview program. A total of 466, 246, 490, and 400 nucleotides from \(\text{vif}, \text{vpu}, \text{env}, \text{and nef}\) genes, respectively, were analyzed and compared to the sequence of SHIV\(_{KU-1bMC33}\) using the SE Central Software package.

**Immunoprecipitation assays**

To determine if macaques developed antibodies to SHIV proteins following inoculation, the plasma obtained at necropsy was used in immunoprecipitation assays. C8166 cells were inoculated with approximately \(10^4\) TCID\(_{50}\) of SHIV\(_{KU-1bMC33}\) for 4 days. The cells were then incubated in methionine/cysteine-free media for 2 h and radiolabeled with 500 µCi of \(^{35}\)S-methionine/cysteine for 15 hours. The cells were lysed in 1 ml of 1x RIPA buffer (50 mM Tris-HCl, pH7.5; 50 mM NaCl, 0.5% deoxycholate; 0.2% SDS; 10 mM EDTA) and nuclei were removed by centrifugation. The cell lysates were incubated overnight at 4°C with 10 µl of plasma from each macaque and protein-A Sepharose beads (Sigma). The immunoprecipitates bound to the beads were washed 3x in 1x RIPA, resuspended in 75 µl of 2x sample reducing buffer (125 mM Tris-HCl pH 6.8; 4% SDS, 10% glycerol. 0.06% bromophenol blue, 1% \(\beta\)-mercaptoethanol), and boiled for 5 min. Proteins were separated on a 10% SDS-PAGE gel and proteins visualized by autoradiography. Controls included pooled prebleed plasma from macaques RAK10, RCS10, and RPL10 (negative control) and plasma from macaques that had been previously inoculated with non-pathogenic SHIV-4 (positive control).
VII. Chapter 3: Replication and persistence of simian-human immunodeficiency viruses in rhesus macaques expressing Vif proteins with mutations in the SLQYLA and HCCH domains

Abstract

The Vif protein of primate lentiviruses interacts with APOBEC3 proteins, which results in shunting of the APOBEC3-Vif complex to the proteosome for degradation. Using the simian-human immunodeficiency virus (SHIV)/macaque model, we compared the replication and pathogenicity of SHIVs that express a Vif protein in which the entire SLQ(Y/F)LA (SHIV\textsubscript{Vif5A}) or HCCH (SHIV\textsubscript{VifHCCH(-)}) domains were substituted with alanine residues. Each virus was inoculated into three macaques and various viral and immunological parameters followed for six months. All macaques maintained stable circulating CD4\textsuperscript{+} T cells, developed low viral loads, maintained the engineered mutations, yielded no histological lesions, and developed immunoprecipitating antibodies early post-inoculation. Sequence analysis of \textit{nef} and \textit{vpu} from three lymphoid tissues revealed a high percentage of G-to-A-substitutions. Our results show that while the presence of HCCH and SLQ(Y/F)LA domains are critical \textit{in vivo}, there may exist APOBEC3 negative reservoirs that allow for low levels of viral replication and persistence but not disease.
Introduction

Mammals have evolved cellular restriction factors such as apolipoprotein B mRNA editing enzyme catalytic peptide-like 3G (APOBEC3G; hA3G) to inhibit the replication of retroviruses. However, retroviruses, such as human immunodeficiency virus type-1 (HIV-1), have also developed mechanisms to counter defend these restriction factors by subverting the host’s proteosomal degradation pathway. For instance, the HIV-1 Vif protein has been shown to interact with A3G promoting its accelerated degradation by the 26S proteosome (Sheehy et al., 2003). APOBEC3G is a cytidine deaminase that, if packaged into HIV-1Δvif/SIVmac239Δvif virions, induces cytidine deamination of newly synthesized minus-strand viral DNA from cytosines to uracils, leading to G to A transitions in plus strand synthesis and inhibition of viral replication (Jarmuz et al., 2002; Harris et al., 2003; Mariani et al., 2003; Mangeat et al., 2003; Sheehy et al., 2003; Virgen et al., 2007; Yu et al., 2004; Zennou et al., 2006).

Previous in vitro studies have shown that SIVmac239Δvif has been shown to be restricted by hA3G, hA3F, hA3H and to a lesser extent hA3B, hA3C, hA3D (Dang et al., 2006; Dang et al., 2008; Mariani et al., 2003; Yu et al., 2004b; Zennou et al., 2006). The HIV-1 Vif has limited activity against rhesus and African green monkey A3 proteins while Vif from SIVmac239 and SIVagm have broader specificities. While less is presently known about the rhesus A3 proteins, it is known that HIV-1Δvif can be inhibited by rhA3G, rhA3F, rhA3B, and to a lesser extent rhA3H and rhA3DE (Virgen et al., 2007). SIVmac239Δvif has been shown to be restricted by rhA3G, rhA3F, rhA3C, rhA3B and rhA3DE, and to a lesser extent rhA3H (Virgen et al., 2007; Zennou et al., 2006).
Sequence analysis of Vif proteins from different lentiviruses reveals that there are two highly conserved domains in the carboxyl terminus of Vif required for the assembly of the Vif-CBF-β-Cul5/ElonginB/C/Rbx1 E3 ubiquitin ligase complex, the BC box

144SLQ(Y/F)LAL150 and Zn++ binding (HCCH; H108-X5-C114-X17-18-C133-X3-5-H139) motifs (Jäger et al., 2011; Kile et al., 2002; Luo et al., 2005; Mehle et al., 2004; Mehle et al., 2006; Paul et al., 2006; Yu et al., 2003; Yu et al., 2003). As described in the Introduction, the 144SLQ(Y/F)LAL150 motif sequence resembles a conserved motif identified in the BC box of the suppressors of cytokine signaling (SOCS) proteins and mediates high affinity binding of Vif to elongin C (Luo et al., 2005; Mehle et al., 2004a; Mehle et al., 2004b; Mehle et al., 2006; Stopak et al., 2003; Yu et al., 2003; Yu et al., 2004c). Located upstream of the BC box is the HCCH motif, which was found to mediate interaction with Cullin-5. Previous studies have shown that substitution of either critical cysteine residue prevents the binding of Zn++ thus affecting the proper folding of the protein (Paul et al., 2006; Xiao et al., 207). Following the binding of N-terminus of Vif to the N-terminus of A3G, both the SLQ(Y/F)LAL and HCCH motifs enable Vif to recruit the Cul5/ElonginB/C/Rbx1 E3 ubiquitin ligase complex to accelerate the polyubiquitination and subsequent 26S proteosomal degradation (Conticello et al., 2003; Marin et al., 2003; Mehle et al., 2004; Stopak et al., 2003; Yu et al., 2003; Yu et al., 2004).

In this study, we constructed simian-human immunodeficiency viruses (SHIVs) in which five amino acids in the SLQYLA motif (SHIV\textsubscript{Vif5A}) and four amino acids of the HCCH (SHIV\textsubscript{VifHCCH(-)}) domain were changed to alanine residues. Our results indicate that rhA3G and rhA3F was efficiently incorporated into both SHIV\textsubscript{Vif5A} and SHIV\textsubscript{VifHCCH(-)}.
virions. However, rhA3F was stable in the presence of wild type and mutant Vif viruses, while only mutant Vif viruses efficiently degraded rhA3G. *In vitro*, rhesus A3G (rhA3G) was more effective at cytidine deamination than rhesus A3F (rhA3F). Macaques inoculated with mutant Vif viruses maintained stable circulating CD4+ T cells, developed low viral loads, and developed immunoprecipitating antibodies. Sequence analysis of both *nef* and *vpu* from the thymus, spleen, and small intestine (ileum) of SHIV\textsubscript{VifHCCH(-)} inoculated macaques displayed a high percentage of G-to-A substitutions. Our results show that both the HCCH and SLQYLA domains were critical to Vif function *in vivo* but that production of viral RNA only persisted in the macaques inoculated with the SHIV\textsubscript{VifHCCH(-)}.

**Results**

**Replication of SHIV\textsubscript{Vif5A} and SHIV\textsubscript{VifHCCH(-)} in A3 positive and negative cell lines**

We performed assays to examine the replication of parental SHIV\textsubscript{KU-2MC4}, SHIV\textsubscript{Vif5A}, SHIV\textsubscript{VifHCCH(-)}, SHIV\textsubscript{VifAAQYLA} and SHIV\textsubscript{VifStop} in hA3G/F positive (C8166) and negative (SupT1) cell lines as well as rhesus PBMC (rhA3G/F). The sequence of the Vif mutants that were analyzed in this study are shown in Figure 28. We included SHIV\textsubscript{VifAAQYLA} in these growth curve studies for comparison. The replication of SHIV\textsubscript{VifAAQYLA} has been previously reported in both tissue culture and rhesus macaques (Schmitt et al., 2009). Cells were inoculated with equivalent amounts (25 ng of p27) of each virus and the levels of p27 (Gag) released into the culture medium were quantified
using a commercial antigen capture assay. All four mutant viruses (SHIV<sub>Vif₅A</sub>, SHIV<sub>VifHCCH(-)</sub>, SHIV<sub>VifAAQYLA</sub> and SHIV<sub>VifStop</sub>) replicated in SupT1 cells to similar levels as parental SHIV<sub>KU-2MC⁴</sub> by day 15 post-inoculation, although the kinetics of replication were slower (Figure 29). However, inoculation of SHIV<sub>Vif₅A</sub>, SHIV<sub>VifHCCH(-)</sub>, SHIV<sub>VifAAQYLA</sub> and SHIV<sub>VifStop</sub> into hA3G/F positive C8166 cell cultures resulted in less than 0.01% of the p27 released compared to parental SHIV<sub>KU-2MC⁴</sub> (Figure 29).

**Both SHIV<sub>Vif₅A</sub> and SHIV<sub>VifHCCH(-)</sub> incorporate rhA3G and rhA3F into virus particles**

The above results suggested that hA3G and hA3F or rhA3G and rhA3F might be incorporated into maturing virus particles resulting in restriction of replication. As we are most interested in an rhA3G and rhA3F, we transfected 293 cells with plasmids expressing an HA-tagged rhA3G or a V5-tagged rhA3F and the complete genomes of either SHIV<sub>Vif₅A</sub>, SHIV<sub>VifHCCH(-)</sub>, SHIV<sub>VifStop</sub>, or parental SHIV<sub>KU-2MC⁴</sub>. At 48 hours post-transfection, the culture medium was collected, clarified, and the virus partially purified and concentrated by ultracentrifugation. The amount of p27 was determined followed by Western blot analysis to detect for the presence or absence of rhA3G or rhA3F. The results shown in Figure 30 indicate that rhA3G was incorporated into SHIV<sub>Vif₅A</sub>, SHIV<sub>VifHCCH(-)</sub>, and SHIV<sub>VifStop</sub> virus particles but was excluded from SHIV<sub>KU-2MC⁴</sub>
Figure 28. The sequence of wild type SIV_{mac230} Vif and the two mutants assessed in this study.
**Zn$^{2+}$ Binding**

**Domain**  
SIV$_{mac}$239

\[ ^{108} \text{LHSTYFCFTAGEVRRAIRGEQLSCCRFPRAHKYQVPSLQYLAKVVSDVRSQGENPTWKO}^{162} \]

**BC Box Domain**

SHIV$_{V_{3+}CC}$ (-)

\[ ^{108} \text{LLASTYFPAFTAGEVRRAIRGEQLLSAARFPRAAKYQVPSLQYLALKVVSDVRSQGENPTWKO}^{162} \]

SHIV$_{V_{3+}5A}$

\[ ^{108} \text{LHSTYFCFTAGEVRRAIRGEQLSCCRFPRAHKYQVPAAAAAALKVVSDVRSQGENPTWKO}^{162} \]
Figure 29. Replication of SHIV
KU-2MC4, SHIV
VifHCCH(-), SHIV
Vif5A, and SHIV
VifStop in C8166 cells (hA3G/F positive), SupT1 cells (hA3G/F negative) and rhesus macaque PBMC. Cells were inoculated with equal amounts of each virus (25 ng of p27) and the levels of p27 in the culture supernatants was assessed at various time points post-inoculation. Panel A. Replication of viruses in C8166 cells. Panel B. Replication of viruses in C8166 cells. Panel C. Replication of viruses in activated rhesus macaque PBMC. (■) SHIV
KU-2MC4, (●) SHIV
VifHCCH(-), (◆) SHIV
Vif5A, and (▼) SHIV
VifStop.
A. C8166

- SHIV<sub>KU-2MC4</sub>
- SHIV<sub>VHCCCH(-)</sub>
- SHIV<sub>VtSB</sub>
- SHIV<sub>VtSBp</sub>

Days Post-Infection

B. SupT1

- SHIV<sub>KU-2MC4</sub>
- SHIV<sub>VHCCCH(-)</sub>
- SHIV<sub>VtSB</sub>
- SHIV<sub>VtSBp</sub>

Days Post-Infection

C. Rhesus PBMC

- SHIV<sub>KU-2MC4</sub>
- SHIV<sub>VHCCCH(-)</sub>
- SHIV<sub>VtSB</sub>
- SHIV<sub>VtSBp</sub>

Days Post-Infection

p27 Released (ng/mL)
Figure 30. The incorporation of rhA3G and rhA3F into virus particles. 293 cells were co-transfected with a plasmid that expresses the genome of either SHIV_{KU-2MC4}, SHIV_{Vif5A}, SHIV_{VifHCCH(-)} or SHIV_{VifStop} and a vector expressing HA-rhA3G or V5-rhA3F. At 48 hours, the culture medium was collected, clarified by low speed centrifugation and concentrated by ultracentrifugation through a 20/60 (w/v) step gradient. Equivalent levels of p27 from each sample were resuspended in sample reducing buffer, boiled and the proteins were separated by 10% SDS-PAGE. The presence of either rhA3G or rhA3F was detected through western blot analysis using an antibody directed by against the HA-tag or V5-tag, respectively. Panel A. Results of the rhA3G incorporation into virus particles. Lane 1. RhA3G proteins detected in lysate from cells transfected with the vector expressing HA-rhA3G alone. Lane 2-5. RhA3G proteins detected in virus from cells transfected with a vector expressing HA-rhA3G and SHIV_{KU-2MC4} (Lane 2), SHIV_{Vif5A} (Lane 3), SHIV_{VifHCCH(-)}, (Lane 4) and SHIV_{VifStop} (Lane 5). Lane 6. RhA3G proteins in pelleted supernatant from cells transfected with empty expression vectors. Panel B. The blot in Panel A was stripped and reprobed with an antibody against the p27 protein. Panel C. Results of the rhA3G incorporation of rhA3F into virus particles. Lane 1. RhA3F proteins detected in a lysate from cell transfected with a vector expressing rhA3F alone. Lane 2-5. RhA3F proteins detected in virus from cells transfected with a vector expressing rhA3F and SHIV_{KU-2MC4} (Lane 2), SHIV_{Vif5A} (Lane 3), SHIV_{VifHCCH(-)}, (Lane 4) and SHIV_{VifStop} (Lane 5). Lane 6. RhA3F proteins in pelleted supernatant from cells transfected with empty expression vectors.
vectors. Panel D. The blot in Panel C was stripped and reprobed with an antibody against the p27 protein.
particles. However, we found that the rhA3F protein was incorporated into all four viruses (Figure 30). These results indicate that rhA3G was selectively incorporated only into the Vif mutants while rhA3F was incorporated into all viruses. The inability of SIV_{mac}239 Vif to degrade and prevent the incorporation of rhA3F has been previously reported (Virgen et al., 2007).

**Rhesus A3F is stable in the presence of SIVmac239 Vif**

Since we observed that rhA3F was incorporated into virus particles, we determined the stability of rhA3F and rhA3G in the presence of the viral genome. 293 cells were co-transfected with vectors expressing rhA3G or rhA3F and the genomes of SHIV_{KU-2MC4}, SHIV_{Vif5A}, SHIV_{VifHCCH(-)}, or SHIV_{VifSTOP}. Our results indicate that in the presence of the SHIV_{KU-2MC4} genome, rhA3G was not stable whereas it was detected in the presence of the SHIV_{Vif5A}, SHIV_{VifHCCH(-)} or SHIV_{VifSTOP} (Figure 31). We also found that rhA3F appeared to be stable in the presence of both SHIV_{KU-2MC4} and also SHIV_{Vif5A}, SHIV_{VifHCCH(-)}, or SHIV_{VifSTOP} genomes. However, it should be noted that we consistently found higher levels of rhA3F in the presence of the SHIV_{VifSTOP} genome, indicating fundamental differences between the targeting site-directed Vif mutants and the absence of the Vif protein. The results obtained with rhA3F correlate well with the above experiment showing that rhA3F was incorporated into virus particles.
Figure 31. Stability of the rhA3G and rhA3F proteins in the presence of the various *vif* mutant viral genomes. Plasmids expressing the genomes of either SHIV\textsubscript{KU-2MC4}, SHIV\textsubscript{Vif5A}, SHIV\textsubscript{VifHCCH(-)} or SHIV\textsubscript{VifStop} were co-transfected into 293 cells with a vector expressing either HA-rhA3G or V5-rhA3F. At 24 hours post-transfection, cells were lysed and proteins were precipitated using methanol. Proteins were separated on a 10% SDS-PAGE gel and probed using an antibody directed against either the HA or V5-tag. All samples were normalized to the amount of β-actin protein. Panel A. Stability of rhA3G in the presence of SHIV\textsubscript{KU-2MC4} (Lane 2), SHIV\textsubscript{VifStop} (Lane 3), SHIV\textsubscript{Vif5A} (Lane 4) or SHIV\textsubscript{VifHCCH(-)} (Lane 5) was assessed. Lanes 1 and 6 were used as positive (HA-rhA3G only, Lane 1) and negative (pcDNA3.1(+) only, Lane 6) controls. Panel B. The stability of rhA3F in the presence of SHIV\textsubscript{KU-2MC4} (Lane 2), SHIV\textsubscript{VifStop} (Lane 3), SHIV\textsubscript{Vif5A} (Lane 4) or SHIV\textsubscript{VifHCCH(-)} (Lane 5) was assessed. Lanes 1 and 6 were used as a negative (pcDNA3.1(+) only, Lane 1) or positive (V5-rhA3F only, Lane 6) controls. Similar results were acquired at 48 hours as well (data not shown). Panel B and C. The blot in Panel A and B were stripped and reprobed with an antibody against β-actin protein.
SHIV<sub>VifHCCH(-)</sub>, SHIV<sub>VifSTOP</sub>, or SHIV<sub>KU-2MC4</sub> and a plasmid expressing either an HA-tagged rhA3G, V5-tagged rhA3F or the empty vector. At 48 hours, the virus containing supernatant was collected, clarified by centrifugation, and the p27 levels quantified. Equivalent levels of p27 from each virus stock were used to titrate infectious virus by a series of 10-fold dilutions onto TZM.bl cells. At 48 hours, the cells were stained for β-galactosidase activity (β-galactosidase is under the control of the LTR and expression of Tat from the incoming virus) and the number of TCID<sub>50</sub> determined. If rhA3G or rhA3F was packaged into the virions, the virus should be hypermutated resulting in the expression of a non-functional Tat. As shown in Figure 32, SHIV<sub>KU-2MC4</sub> replicated to approximately the same levels in the presence or absence (empty vector) of rhA3G while the titers of SHIV<sub>VifHCCH(-)</sub>, SHIV<sub>Vif5A</sub> and SHIV<sub>VifSTOP</sub> were 100 to 1000 fold less than the unmodified virus (SHIV<sub>KU-2MC4</sub>). With rhA3F, our results indicate that SHIV<sub>KU-2MC4</sub>, SHIV<sub>VifHCCH(-)</sub> and SHIV<sub>Vif5A</sub> had an approximate 10-fold decrease in titers and SHIV<sub>VifSTOP</sub> had greater than a 1000-fold decrease in titer (Figure 32). The results with rhA3F also emphasize the differences in site-directed mutations in SHIV<sub>VifHCCH(-)</sub>, SHIV<sub>Vif5A</sub> and the severely truncated Vif in SHIV<sub>VifSTOP</sub>.

**Rhesus A3G but not rhA3F causes significant G-to-A mutations in the nef gene of the SHIV genome**

The results above indicated that significant levels of rhA3F were incorporated into virions and rhA3F was stable in cells expressing the wild type genome. We compared the level of G-to-A mutations of the SHIV genome in the presence of either rhA3G or rhA3F. The results are shown in Figure 33 and Table V. Using viral genomes to express the Vif protein, we observed minimal G-to-A changes in SHIV<sub>KU-2MC4</sub> nef in
Figure 32. The incorporation of rhA3G results in less infectious virus. 293 cells were co-transfected with full-length vectors expressing SHIV$_{KU-2MC4}$, SHIV$_{VifHCCH(-)}$, SHIV$_{Vif5A}$ and SHIV$_{VifStop}$ and plasmids expressing either rhA3G of rhA3F. As a control each full-length genome was co-transfected with the empty vector, pcDNA3.1(+), only. At 48 hours, the culture medium was harvested, assessed for p27, and the amount of infectious virus titrated on TZM.bl cells as described in the Material and Methods section. Shown are the infectious titers in the absence or presence of either rhA3G (Panel A) or rhA3F (Panel B).
Table V. Results of the hypermutation assay of *nef* amplified by rhA3G and rhA3F in the presence of wild type and mutant Vif proteins.
<table>
<thead>
<tr>
<th>APOBEC3 protein</th>
<th>Virus</th>
<th># Total number of bases sequenced</th>
<th>Total # of G-to-A mutations/percent G-to-A mutations of bases sequenced</th>
<th>Number and context of cytidine deamination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5'-GA</td>
</tr>
<tr>
<td>rhA3G</td>
<td>SHIV_0103-3MC4</td>
<td>4875</td>
<td>2 (0.04)</td>
<td>1</td>
</tr>
<tr>
<td>rhA3G</td>
<td>SHIV_7</td>
<td>4875</td>
<td>24 (0.49)</td>
<td>23</td>
</tr>
<tr>
<td>rhA3G</td>
<td>SHIV_0103-3MC5</td>
<td>4875</td>
<td>26 (0.53)</td>
<td>21</td>
</tr>
<tr>
<td>rhA3G</td>
<td>SHIV_0103-3MC6</td>
<td>4875</td>
<td>29 (0.59)</td>
<td>27</td>
</tr>
<tr>
<td>rhA3F</td>
<td>SHIV_0103-3MC4</td>
<td>4875</td>
<td>2 (0.04)</td>
<td>1</td>
</tr>
<tr>
<td>rhA3F</td>
<td>SHIV_7</td>
<td>4875</td>
<td>1 (0.02)</td>
<td>1</td>
</tr>
<tr>
<td>rhA3F</td>
<td>SHIV_0103-3MC5</td>
<td>4875</td>
<td>2 (0.04)</td>
<td>1</td>
</tr>
<tr>
<td>rhA3F</td>
<td>SHIV_0103-3MC6</td>
<td>4875</td>
<td>3 (0.06)</td>
<td>2</td>
</tr>
</tbody>
</table>

* A 325 base fragment corresponding to the 5' end of nef was amplified, cloned and 15 individual clones were selected for sequence analysis.
Figure 33. RhA3G but not rhA3F induced G-to-A substitutions in single colony sequences obtained from SIV\textsubscript{mac239} nef. Full-length genomes of wild type and \textit{vif} mutant viruses were co-transfected with vectors expressing either rhA3G or rhA3F into 293 cells. At 48 hours post-transfection, the culture medium was collected, clarified by low speed centrifugation, DNase I-treated and used to infect TZM.bl cells. A 325 base pair fragment of SIV\textsubscript{mac239} nef was amplified from DNA at 24 hours post-infection and assessed for G-to-A substitutions. Panel A. A graph depicting the cumulative number of mutations from the 15 independent clones per virus assessed in the presence of rhA3G. Each bar is shaded according to the proportion of G-to-A substitutions that occurred in the context of GA (red), GG (yellow), GT (cyan), GC (gray), or non G-to-A (black). Panel B. A graph depicting the cumulative number of mutations from the 15 independent clones per virus assessed in the presence of rhA3G. Each bar is shaded according to the proportion of G-to-A substitutions that occurred in the context of GA (red), GG (yellow), GT (cyan), GC (gray), or non G-to-A (black). Panel C. A graph depicting the cumulative number of mutations from the 15 independent clones per virus assessed in the presence of rhA3F.
A. 

Number of G-to-A Substitutions

- Non G-to-A
- GA-to-AA
- GG-to-AG
- GT-to-AT
- GC-to-AC

SHIV_{KU-2MC4}  SHIV_{VifHCCH(-)}  SHIV_{Vif5A}  SHIV_{VifStop}

B. 

SHIV_{KU-2MC4}

SHIV_{VifStop}

SHIV_{VifHCCH(-)}

SHIV_{Vif5A}

Scale: 0 to 325
C.

![Bar chart with different categories and labels](chart.png)

- Non G-to-A
- GA-to-AA
- GG-to-AA
- GT-to-AT
- GC-to-AC

Number of G-to-A Substitutions

SHIV\textsubscript{KU-2MC4} \quad SHIV\textsubscript{VifHCCH(-)} \quad SHIV\textsubscript{Vif5A} \quad SHIV\textsubscript{VifStop}
the presence of either rhA3G or rhA3F. In the presence of rhA3G, we found that SHIV\textsubscript{Vif5A}, SHIV\textsubscript{VifHCCH(-)} or SHIV\textsubscript{VifSTOP} \textit{nef} had an increase in the number of G-to-A changes (from 2 to 24-29) or approximately 0.5% of the bases sequenced (Table V). However, in the presence of rhA3F, we observed that SHIV\textsubscript{Vif5A}, SHIV\textsubscript{VifHCCH(-)}, and SHIV\textsubscript{VifSTOP} \textit{nef} had a similar level of G-to-A changes compared to SHIV\textsubscript{KU-2MC4} (Table V). The results with the rhA3F suggest that incorporation of rhA3F does not result in significant G-to-A substitutions in \textit{nef}.

**Assessment of SHIV\textsubscript{VifHCCH(-)} replication in macaques**

We inoculated three macaques with $10^4$ TCID\textsubscript{50} of SHIV\textsubscript{VifHCCH(-)}. Prior to inoculation, macaques AS34, AS51, and AV18 had circulating CD4\textsuperscript{+} cell levels of 1,509, 1,109, and 400 cells per $\mu$l, respectively. These three macaques maintained levels of circulating CD4\textsuperscript{+} T cells at pre-inoculation levels or higher throughout the course of the six-month infection (Figure 34). All three macaques were euthanized at 26-28 weeks post-inoculation in excellent condition. At necropsy, AS34, AS51, and AV18 had circulating CD4\textsuperscript{+} T cell levels of 2,124, 1,338, and 1,298 cells/$\mu$l, respectively. This contrasts with macaques inoculated with parental SHIV\textsubscript{KU-2MC4}, which developed a severe loss of circulating CD4\textsuperscript{+} T cells within 3-4 weeks post-inoculation (Figure 34). Analysis of the plasma viral loads revealed a mean viral load of $4.15 \times 10^3$ copies per ml at one week post-inoculation (Figure 35), which is approximately 10,000-fold less than the macaques inoculated with parental SHIV\textsubscript{KU-2MC4} (Figure 35). Following the first week of infection, the plasma viral loads in macaques AS34, AS51, and AV18 were at or just above the limits of detection. At necropsy, viral loads in all three macaques were
Figure 34. Circulating CD4$^+$ T cell levels in macaques inoculated with SHIV$_{VifHCCH(-)}$, SHIV$_{Vif5A}$, and SHIV$_{KU-2MC4}$. Panel A. The levels of circulating CD4$^+$ T cells in three macaques (AS34, ●; AS51, ■; and AV18, ▲) following inoculation with SHIV$_{VifHCCH(-)}$. Panel B. The levels of circulating CD4$^+$ T cells in three macaques (CX54, ●; ER65, ■; and I92, ▲) following inoculation with SHIV$_{Vif5A}$. Panel C. The levels of circulating CD4$^+$ T cells in three macaques (44O, ●; 44N, ■; 35N, ▲) following inoculation with SHIV$_{KU-2MC4}$. 
Figure 35. Plasma viral loads in macaques inoculated with SHIV\textsubscript{VifHCCH(-)}, SHIV\textsubscript{Vif5A}, and SHIV\textsubscript{KU-2MC4} as determined by real-time quantitative PCR. Panel A. Plasma viral loads in three macaques (AS34, ●; AS51, ■; and AV18, ▲) following inoculation with SHIV\textsubscript{VifHCCH(-)}. Panel B. Plasma viral loads in three macaques (CX54, ●; ER65, ■; and I92, ▲) following inoculation with SHIV\textsubscript{Vif5A}. Panel C. Plasma viral loads in three macaques (44O, ●; 44N, ■; 35N, ▲) following inoculation with SHIV\textsubscript{KU-2MC4}.
undetectable. These results indicate that SHIV<sub>VifHCCH(-)</sub> transiently replicated in these macaques prior to control of the infection.

**Assessment of SHIV<sub>Vif5A</sub> replication in macaques**

We also assessed the ability of SHIV<sub>Vif5A</sub> to replicate in three macaques (CX54, ER65, and I92). Prior to inoculation, macaques CX54, ER65, and I92 had circulating CD4<sup>+</sup> T cell levels of 777, 844, and 1,451 cells per µl, respectively (Figure 34). These macaques, similar to those inoculated with SHIV<sub>VifHCCH(-)</sub>, maintained circulating CD4<sup>+</sup> T cell levels near the pre-inoculation levels throughout the course of the 6 month infection. All three macaques were euthanized at 26 weeks post-inoculation in excellent condition. At necropsy, CX54, ER65, and I92 had circulating CD4<sup>+</sup> T cell levels of 513, 853, and 1192 cells per µl, respectively (Figure 34). This also contrasts with macaques inoculated with parental SHIV<sub>KU-2MC4</sub> (Figure 34). Analysis of the plasma viral loads of the macaques inoculated with SHIV<sub>Vif5A</sub> revealed a mean viral load of 1.3 x10<sup>3</sup> copies per ml at one week post-inoculation (Figure 35). Following the first week of infection, the plasma viral loads in macaques CX54, ER65, and I92 fell to undetectable levels.

**Mutations in vif amplified from isolated DNA of PBMC during the course of infection**

We assessed the stability of the engineered mutations in the vif gene from DNA isolated PBMC throughout the course of infection. The vif sequence was amplified, directly sequenced and compared to the input vif sequence of SHIV<sub>Vif5A</sub> or SHIV<sub>VifHCCH(-)</sub>. We observed that the engineered mutations were stable throughout the course of the six month infection for macaques inoculated with SHIV<sub>VifHCCH(-)</sub>. We performed bulk
Figure 36. Analysis of *vif* sequences amplified from DNA isolated from PBMC at various times post-inoculation. Panel A. Quantitative representation of mutations induced in bulk sequences obtained from a 450 base pair fragment of SHIV<sub>VifHCCH(−)</sub> Vif. Each vertical bar represents the total number of G-to-A substitutions. Each bar is shaded according to the proportion of G-to-A substitutions that occurred in the context of GA (red), GG (yellow), GT (cyan), or non G-to-A (black). Panel B. A 450 base pair fragment of the SIV Vif was amplified using DNA isolated from macaque PBMC inoculated with SHIV<sub>VifHCCH(−)</sub>. Mutations from bulk sequence analysis from each macaque are shown. Each mutation is denoted by a vertical line that is color coded with respect to the dinucleotide context: GA (red), GG (yellow), GT (cyan), GC (gray) and non-G-to-A substitutions (black).
sequencing of the \textit{vif} amplified at 1, 2, 3, 4, 6, and 16 weeks. As shown in Figure 36, the number of G-to-A mutations in the amplified \textit{vif} gene increased with time in two (AS51, AV18) of the three macaques while the third macaque (AS34) appeared to have the largest number of G-to-A mutations at 3 weeks post-inoculation. As we had reported earlier, the majority of the G-to-A mutations were in the context of the 5'-TC rather than 5'-CC \citep{Schmitt09}. We were able to amplify part of the \textit{vif} sequence at 1 week from PBMC DNA isolated from macaques inoculated with SHIV\textsubscript{Vif5A} (CX54, ER65, I92). The sequence analysis also revealed that these \textit{vif} mutations were stable at one week post-inoculation (data not shown). However, from 4 weeks until necropsy, we were unable to amplify the \textit{vif} gene, which correlates with the undetectable plasma viral loads in these macaques after week 1 post-inoculation.

\textbf{Macaques inoculated with SHIV\textsubscript{Vif5A} and SHIV\textsubscript{VifHCCH(-)} developed anti-viral antibody responses early after inoculation}

We analyzed the plasma from infected macaques at 12 weeks and at necropsy (26-28 weeks) for the presence of immunoprecipitating antibody responses. At 12 weeks post-inoculation, all three macaques inoculated with SHIV\textsubscript{Vif5A} had developed antibody responses to p27 but only macaque ER65 developed antibody responses to the Env glycoprotein (Figure 37). For macaques inoculated with SHIV\textsubscript{VifHCCH(-)}, all three macaques developed antibodies to p27 and the Env glycoprotein (Figure 37). At necropsy, we could not detect the presence of immunoprecipitating antibodies from macaques inoculated with either SHIV\textsubscript{Vif5A} or SHIV\textsubscript{VifHCCH(-)} (Figure 37). A macaque inoculated with parental SHIV\textsubscript{KU-2MC4} (44O) did not develop antibodies to the virus at either time point, which is common for macaques that develop severe CD4\textsuperscript{+} T cell loss.
Figure 37. Macaques inoculated with SHIV\textsubscript{VifHCCH(−)} and SHIV\textsubscript{VifSA} developed antibody responses against each virus early after inoculation. C8166 cells were inoculated with SHIV\textsubscript{KU-2MC4} for 5 days, starved in methionine/cysteine-free DMEM, and then radio labeled overnight with \textsuperscript{35}S-methionine/cysteine. The culture medium was harvested and used in immunoprecipitation reactions with plasma from AS34, AS51, AV18, CX54, ER65, and I92 as described in the Materials and Methods section. The immunoprecipitates were washed with 1 x RIPA buffer and boiled in 2x sample reducing buffer. The proteins were then separated on a 12%-SDS-PAGE gel and visualized using autoradiographic techniques. Panel A. SHIV proteins immunoprecipitated with plasma obtained at week 12 post-inoculation. Panel B. SHIV proteins immunoprecipitated with plasma obtained at necropsy. (Lanes 1-3) SHIV\textsubscript{VifHCCH(−)} proteins immunoprecipitated using plasma from either AS34 (Lane 1), AS51 (Lane 2), or AV18 (Lane 3), respectively. (Lanes 4-6) SHIV\textsubscript{Vif5A} proteins immunoprecipitated using plasma from either CX54 (Lane 4), ER65 (Lane 5), or I92 (Lane 6), respectively. (Lane 7) SHIV proteins immunoprecipitated using plasma from an uninfected macaque. (Lane 8) SHIV proteins immunoprecipitated from a positive control plasma sample. The molecular weight standards are shown on the left.
during the acute phase (<4 weeks) following inoculation with a pathogenic X4 SHIV (data not shown).

**Histological examination of tissues and viral loads in visceral organs of macaques inoculated with SHIV\textsubscript{Vif5A} and SHIV\textsubscript{VifHCCH(-)}**

Tissues from infected macaques were examined for the presence of lesions consistent with this pathogenic X4 SHIV (Joag et al., 1998). Macaques inoculated with either SHIV\textsubscript{Vif5A} or SHIV\textsubscript{VifHCCH(-)} did not exhibit histological lesions in any of the 13 visceral organs examined or in the CNS (data not shown). As the plasma viral loads were significantly less than macaques inoculated with SHIV\textsubscript{KU-2MC4}, we assessed the tissue distribution of these two viruses in macaques. RNA was isolated from 13 visceral organs from macaques inoculated with SHIV\textsubscript{Vif5A} and SHIV\textsubscript{VifHCCH(-)}, DNase I treated, and copy numbers determined by real time quantitative RT-PCR. The number of copies is expressed per $10^6$ copies of GAPDH. The results obtained from macaques inoculated with SHIV\textsubscript{VifHCCH(-)} are shown in Figure 38. We found that macaque AS34 had eight tissues (kidney, lung, pancreas, axillary lymph node, inguinal lymph node, spleen, thymus and tonsil) greater than 1,000 copies per $10^6$ copies of GAPDH (Figure 38). Macaque AS51 was found to have two tissues (lung and small intestine) with greater than 1,000 copies per $10^6$ copies of GAPDH, and macaque AV18 had six tissues (lung, pancreas, axillary lymph node, inguinal lymph node, mesenteric lymph node and spleen) with greater than 1,000 copies per $10^6$ copies of GAPDH (Figure 38). Interestingly, the lung had the highest viral copy number in each macaque. In contrast, we were unable to quantify viral RNA from tissues of macaques that were inoculated
Figure 38. Comparison of viral copy numbers in the visceral tissues of macaques inoculated with SHIV\textsubscript{KU-2MC4}, SHIV\textsubscript{VifHCCH(-)}, and SHIV\textsubscript{Vif5A}. RNA was isolated from different visceral organs of macaques inoculated with either SHIV\textsubscript{VifHCCH(-)} or SHIV\textsubscript{Vif5A}. The RNA was DNase I treated to remove residual DNA and used in real time quantitative RT-PCR using oligonucleotide primers specific for SIV gag as described in the Materials and Methods. The levels of viral RNA are shown per 106 copies of GAPDH RNA. Panel A. The results of the real time quantitative RT-PCR using RNA isolated from tissues of macaque 44O, which was inoculated with SHIV\textsubscript{KU-2MC4}. Panels B-D. The results of the real time quantitative RT-PCR using RNA isolated from the tissues of macaques AS34 (Panel B), AS51 (Panel C), and AV18 (Panel D), which were inoculated with SHIV\textsubscript{VifHCCH(-)}. Panels E-G. The results of the real time quantitative RT-PCR using RNA isolated from tissues of macaques CX54 (Panel E), ER65 (Panel F), and I92 (Panel G), which were inoculated with SHIV\textsubscript{Vif5A}.
E.

F.

G.

Viral Copies per 10^6 GADPH copies

Heart  Kidney  Liver  Lung  Sal. Gland  Pancreas  Axillary LN  Mesenteric LN  Sm. Intestine  Spleen  Thymus  Tonsil  PBMC

Viral Copies per 10^6 GADPH copies

Heart  Kidney  Liver  Lung  Sal. Gland  Pancreas  Axillary LN  Mesenteric LN  Sm. Intestine  Spleen  Thymus  Tonsil  PBMC

Viral Copies per 10^6 GADPHcopies

Heart  Kidney  Liver  Lung  Sal. Gland  Pancreas  Axillary LN  Mesenteric LN  Sm. Intestine  Spleen  Thymus  Tonsil  PBMC

207
with SHIV<sub>Vif5A</sub> (data not shown). Macaques inoculated with parental SHIV<sub>KU-2MC4</sub>, 44O and 44N, had high copy numbers of virus in various tissues (Figure 38). Taken together, these results indicate that both SHIV<sub>VifHCCH(-)</sub> and SHIV<sub>Vif5A</sub> infections were significantly less widespread compared to macaques inoculated with parental SHIV<sub>KU-2MC4</sub>.

**G-to-A mutations in nef and vpu amplified from DNA of primary and secondary lymph organs**

A well documented feature of HIV-1Δvif infection is the incorporation of select APOBEC3 proteins into viral particles leading to cytosine deamination during minus strand DNA synthesis (Yu et al., 2003). This results in hypermutation (guanosine to adenosine transitions) of the viral genome and inhibition of virus replication. We examined the number of G-to-A mutations in the nef and vpu genes amplified from DNA isolated from one primary lymphoid organ (thymus) and two secondary lymphoid organs (ileum of the small intestine and spleen) from the three macaques inoculated with SHIV<sub>VifHCCH(-)</sub> and one macaque inoculated with SHIV<sub>KU-2MC4</sub>. We first PCR amplified nef and vpu sequences from tissue DNA and directly sequenced the amplified products. The results showed that there were multiple G-to-A mutations in the nef and vpu regions amplified from three tissues. The number of G-to-A mutations ranged from 54 to 88 or 1.2 to 1.9 percent of the total base pairs sequenced (Figure 39; Table VI). While the percentage of mutations observed in the vpu was lower (perhaps reflecting the gradient of mutations from nef to vpu), the overall number of G-to-A mutations in the context of 5'-TC to 5'-CC was approximately 2:1 (75 versus 31 mutations) (Figure 39; Table VI).
Figure 39. Quantitative assessment of mutations in the nef gene. Panels A-D. A 300 base pair fragment of nef was amplified using DNA isolated from the small intestine (ileum), spleen, and thymus of macaques inoculated with SHIVVifHCCCH(-) and one macaque inoculated with SHIVKU-2MC4. Mutations obtained using sequence analysis of fifteen independent clones from each macaque is shown. Each vertical bar represents an individual clone, whose height represents the total number of substitutions. Each bar is shaded according to the proportion of G-to-A substitutions that occurred in the dinucleotide context: non-G-to-A (black), GA (red), GG (yellow), GT (cyan), or GC (gray). Panels E-G. A representative hypermutation plot from each macaque illustrating the G-to-A and non-G-to-A substitutions in the clones from the ileum amplified DNA. Mutations obtained from bulk (top line) sequence analysis of fifteen independent clones from each macaque are shown. Each mutation is denoted by a vertical line that is color coded with respect to the the dinucleotide context: non-G-to-A (black), GA (red), GG (yellow), GT (cyan), or GC (gray).
Table VI. Results of sequence analysis of nef amplified from the thymus, ileum, and spleen.
<table>
<thead>
<tr>
<th>Animal Tissue</th>
<th>Total number of bases sequenced</th>
<th>Total # of G-to-A mutations/percent G-to-A mutations of bases sequenced</th>
<th>Number and context of cytidine deamination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5'-GA</td>
<td>5'-GG</td>
</tr>
<tr>
<td>AS34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thymus</td>
<td>4500</td>
<td>71/1.6</td>
<td></td>
</tr>
<tr>
<td>Spleen</td>
<td>4500</td>
<td>61/1.4</td>
<td></td>
</tr>
<tr>
<td>Ileum</td>
<td>4500</td>
<td>83/1.8</td>
<td></td>
</tr>
<tr>
<td>AS51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thymus</td>
<td>4500</td>
<td>75/1.7</td>
<td></td>
</tr>
<tr>
<td>Spleen</td>
<td>4500</td>
<td>62/1.4</td>
<td></td>
</tr>
<tr>
<td>Ileum</td>
<td>4500</td>
<td>66/1.5</td>
<td></td>
</tr>
<tr>
<td>AV18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thymus</td>
<td>4500</td>
<td>54/1.2</td>
<td></td>
</tr>
<tr>
<td>Spleen</td>
<td>4500</td>
<td>58/1.3</td>
<td></td>
</tr>
<tr>
<td>Ileum</td>
<td>4500</td>
<td>88/2.0</td>
<td></td>
</tr>
</tbody>
</table>

*a A 300 base fragment corresponding to the 5' end of nef was amplified, cloned and 15 separate clones sequenced.
Figure 40. Quantitative assessment of mutations in the vpu gene. Panels A-D. A 184 base pair fragment of vpu was amplified using DNA isolated from the small intestine (ileum), spleen, and thymus of macaques inoculated with SHIV\textsubscript{VifHCCH(−)} and one macaque inoculated with SHIV\textsubscript{KU-2MC4}. Substitutions were determined using sequence analysis of fifteen independent clones from each macaque are shown. Each vertical bar represents an individual clone, whose height represents the total number of substitutions. Each bar is shaded according to the proportion of G-to-A substitutions that occurred in the dinucleotide context: non-G-to-A (black), GA (red), GG (yellow), GT (cyan), or GC (gray). Panels E-G. A representative hypermutation plot from each macaque illustrating the G-to-A and non-G-to-A substitutions in the clones from the ileum amplified DNA. Mutations obtained from bulk (top line) sequence analysis of fifteen independent clones from each macaque are shown. Each mutation is denoted by a vertical line that is color coded with respect to the the dinucleotide context: non-G-to-A (black), GA (red), GG (yellow), GT (cyan), or GC (gray).
Table VII. Results of sequence analysis of *vpu* amplified from the thymus, ileum, and spleen.
<table>
<thead>
<tr>
<th>Animal Tissue</th>
<th>Total number of bases sequenced</th>
<th>Total # of G-to-A mutations/percent G-to-A mutations of bases sequenced</th>
<th>Number and context of cytidine deamination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5'-GA</td>
<td>5'-GG</td>
</tr>
<tr>
<td>AS34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thymus</td>
<td>2760</td>
<td>13/0.47</td>
<td>6</td>
</tr>
<tr>
<td>Spleen</td>
<td>2760</td>
<td>15/0.54</td>
<td>10</td>
</tr>
<tr>
<td>Ileum</td>
<td>2760</td>
<td>15/0.54</td>
<td>8</td>
</tr>
<tr>
<td>AS51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thymus</td>
<td>2760</td>
<td>14/0.51</td>
<td>11</td>
</tr>
<tr>
<td>Spleen</td>
<td>2760</td>
<td>9/0.33</td>
<td>7</td>
</tr>
<tr>
<td>Ileum</td>
<td>2760</td>
<td>15/0.54</td>
<td>9</td>
</tr>
<tr>
<td>AV18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thymus</td>
<td>2760</td>
<td>3/0.12</td>
<td>3</td>
</tr>
<tr>
<td>Spleen</td>
<td>2760</td>
<td>30/1.09</td>
<td>16</td>
</tr>
<tr>
<td>Ileum</td>
<td>2760</td>
<td>11/0.40</td>
<td>5</td>
</tr>
</tbody>
</table>

* A 184 base fragment corresponding to vpu was amplified, sub-cloned and 15 independent clones sequenced.
We confirmed the presence of these G-to-A mutations by sequencing 15 individual clones of the *nef* and *vpu* genes. We plotted the G-to-A mutations obtained from 15 single colony sequences of *nef* and *vpu* amplified from DNA isolated from the ileum of macaques inoculated SHIV<sub>VifHCCH(-)</sub> (Figures 39 and 40). These results showed that many of the G-to-A mutations found in the direct bulk sequencing products were also found in the individually cloned both genes (Figures 39 and 40). Similar to what we reported previously, the majority of the changes were in the context of 5′-TC (minus strand) (Table VI and VII). Macaque 44O, (inoculated with SHIV<sub>KU2-MC4</sub>) had very few G-to-A substitutions in *vpu* and *nef* amplified from the same three lymphoid organs, which is similar to other macaques inoculated with this parental virus (unpublished observations). Macaque AV18 yielded the highest number of G-to-A substitutions (88) in the ileum of the small intestine. Taken together these findings suggest that one or more rhA3 proteins that have target sequence preference for 5′-TC is the major cytidine deaminase at work *in vivo*.

**Discussion**

There are only a limited number of studies that have analyzed the role of the Vif protein in the SIV or SHIV/macaque models of HIV-1 disease. In the first study, the role of various accessory proteins expressed from SIV<sub>mac239</sub> were evaluated in terms of potential to cause disease (Desrosiers et al., 1998). These investigators found that an SIV with deletion of *vif* had only transient viral loads, could only be detected in PBMC for
a short period of time and the macaques did not produce antibodies against SIV proteins. In a second study, investigators showed that a vif-deleted virus would generate antibody and cellular immune responses if the virus was administered 2-3 times but was not protective against challenge (Sparger et al., 2008). In both studies, the lymphoid and non-lymphoid tissues were not examined for the presence of viral RNA from these vif-deleted viruses. Finally, we previously showed that the introduction of amino acid substitutions into the SLQYLA domain of Vif (to AAQYLA) resulted in a virus (SHIV_{VifAAQYLA}) that did not replicate well in hA3G/F positive cells. However, following inoculation into macaques, the plasma viral loads were lowered from approximately $10^7$ copies per ml for wild type virus to approximately $10^5$ copies per ml for the SHIV_{VifAAQYLA} during the acute phase of infection (Schmitt et al., 2009). In these macaques, the levels of circulating CD4$^+$ T cells were maintained near pre-inoculation levels and they did not develop histological lesions consistent with a pathogenic X4 SHIV. Sequence analysis of the vif gene amplified from the PBMC revealed that the engineered AAQYLA motif had mutated to TAQYLA in two of three macaques. Since serine and threonine are hydroxylated amino acids that differ by a methyl group, it was possible that these Vif proteins may have been “partially” functional in vivo to permit limited spread within macaques. Examination of the tissues from these macaques revealed that viral RNA was present in select tissues, and the viral sequences contained low levels of G-to-A substitutions. We also showed that the inoculation of plasma from two of these macaques into a naive macaque did not result in detectable viral loads up to five months post-inoculation (Schmitt et al., 2009). Thus, despite the low level of viral replication in
cell culture systems, SHIV\textsubscript{VifAAQYLA} was still capable of significant \textit{in vivo} replication, thus validating the need to examine targeted Vif mutants in macaques.

In the present study, we have extended our \textit{in vivo} studies using SHIV mutants that express either a Vif protein with the first five amino acids of the SLQYLA domain substituted with alanines (SHIV\textsubscript{Vif5A}) or a Vif with the HCCH domain substituted with alanines (SHIV\textsubscript{VifHCCH(-)}). Interestingly, our restriction assays showed that incorporation of rhA3G but not rhA3F resulted in the greatest reduction of replication, which is somewhat different from the results previously reported (Zennou and Bieniasz, 2006). Our results indicate that unlike the parental virus, both viruses incorporated rhA3G while both the parental and mutant viruses incorporated rhA3F into virions. Our analysis of rhA3F and rhA3G stability in the presence of viral genome indicated that rhA3F is stable while rhA3G is degraded. The incorporation of rhA3F into wild type virions is in agreement with a previous study (Zennou and Bieniasz, 2006; Virgen and Hatzioannou, 2007). These paradoxical findings suggest that rhA3F can be incorporated in the presence of SIV\textsubscript{mac} Vif and that the virus can seemly replicate without the accumulation of lethal mutations. In the former study, these investigators found that rhA3F was resistant to the SIV\textsubscript{mac} Vif protein. They suggested that rhA3F may not exert a strong negative presence on SIV\textsubscript{mac} \textit{in vivo} and that the SIV\textsubscript{mac} Vif may actually permit some degree of mutagenesis by rhA3F. This was supported by analysis of HIV-1 \textit{gag}, \textit{pol}, and \textit{env} for potential editing sites (GG or GA) that would yield a termination codon if mutated. Their analysis revealed that the percentage of GA editing sites leading to termination codons was lower than GG sites in the \textit{gag}, \textit{pol}, and \textit{env}. Analysis of the SIV \textit{gag}, \textit{pol}, \textit{vif}, \textit{vpr}, \textit{vpx}, and \textit{nef} (the SIV genes of our SHIV) also revealed this trend.
(unpublished observations) yet the mutations observed in the SIV\textsubscript{mac} nef were predominantly in the context of GA and not GG. Whether this was due to lethal GG-to-AG mutations that were subsequently cleared (or not detectable) or whether this was due to rhA3F or another rhA3 protein is unknown at this time. However, it should be noted that there were differences in our study and the previous study. First, in the previous study the investigators used a HIV-1\textsubscript{Δvif} genome and we used an SIV-based viral genome with targeted mutations in the \textit{vif}. We showed that replication of SHIV\textit{VifSTOP} was much lower compared to SHIV\textit{Vif5A} and SHIV\textit{VifHCCH(-)} in the presence of rhA3F. Thus, the results of SHIV\textit{VifSTOP} in our study and HIV-1\textsubscript{Δvif} in their study is comparable. Our study also shows that our viruses with targeted mutations (SHIV\textit{Vif5A} and SHIV\textit{VifHCCH(-)}) were restricted less efficiently in the presence of rhA3F compared to SHIV\textit{VifSTOP}, suggesting that the presence of an intact, “crippled” Vif protein can impose some restriction compared to having a \textsubscript{ΔVif} virus. This may relate to the ability of the Vif\textsubscript{HCCH(-)} and the Vif\textsubscript{5A} proteins to still bind to rhA3F. Previous studies have defined regions of the HIV-1 Vif that are required for binding various hA3 proteins. Many of these regions are in the amino terminal region of the protein. Previous studies demonstrated that HIV-1 Vif residues \textsubscript{40}YRHHY\textsuperscript{44} and \textsubscript{12}QVDRMR\textsuperscript{17} are important for the interaction with hA3G and hA3F/C, respectively (Russell and Pathak, 2007; Mehle et al., 2007). In a more recent study, a conserved YXXL domain at residues 69-72 of the HIV-1 Vif was found to be important for the interaction of Vif with hA3G, hA3F and hA3C (Pery et al., 2009). While the \textsubscript{12}QVDRMR\textsuperscript{17} and \textsubscript{40}YRHHY\textsuperscript{44} are not conserved in the SIV\textsubscript{mac239} Vif, there is a lysine at position 27 of the SIV\textsubscript{mac} Vif that is required for interaction with rhA3G (Chen et al., 2009; Dang et al., 2009). The YWGL domain of the
HIV-1 Vif is also conserved in the SIV Vif (as YWHL). Thus, it is possible that the Vif proteins with targeted amino acid substitutions may still bind and incorporate rhA3F into virus particles but may be inefficient in deamination because of its association with Vif. A second difference in the two studies is the hypermutation assay used to analyze the percentage of G-to-A changes in the genome. While the previous study showed that rhA3F caused G-to-A hypermutation at levels higher than rhA3G, we found that rhA3F produced little hypermutation in the viral genome. In the previous study, these investigators used a VSV pseudotyped HIV-1 vector with the vif deleted as well as other genes and analyzed hrGFP sequences. In our system, we used complete SIV viral genomes expressing the Vif mutant.

One goal of these studies was to determine if SHIV_{Vif5A} was effectively controlled by macaques during the primary phase of infection or if viral variants would emerge that would permit limited replication in macaque tissues. Our results indicate that following inoculation of SHIV_{Vif5A} into macaques, there was a further reduction of plasma viral loads compared to SHIV_{VifAAQYA} (Schmitt et al., 2009). The mean peak plasma virus load (week 1) was 1.3 x 10^3 viral copies per ml or approximately 4,000-fold lower than parental SHIV_{KU-2MC4}. However, following this initial burst of replication, viral RNA from macaques inoculated with SHIV_{Vif5A} was near or below the limits in this assay (~180 copies). It should be noted that the virus was not completely eliminated because viral RNA was occasionally detected in the plasma throughout the 6-month infection using nested RT-PCR (data not shown). At necropsy, the number of copies of viral RNA could not be quantified from various organs of these macaques (detection limit 180 copies) but could be detected by nested RT-PCR. We were unable to detect the virus in PBMC by 3
weeks post-inoculation and were unable to amplify the nef gene at necropsy for sequence analysis. While the results presented here may be predicted based on studies in cell culture systems, these results demonstrate for the first time that a primate lentivirus expressing a Vif protein with only targeted mutations in a critical domain can be completely controlled by the host. Taken together, these results indicate that the SHIV\textsubscript{Vif5A} replicated similarly to a virus that does not express a functional Vif protein (Desrosiers et al., 1998).

No studies have analyzed the role of the Zn\textsuperscript{++} binding domain of Vif \textit{in vivo}. Previous studies have shown that this domain is critical to Vif function, interacting with Cul 5 of the Cul 5/ ElonginB/C/ Rbx E3 ligase (Luo et al., 2005; Xiao et al., 2007). Our results showed that these macaques had slightly higher initial viral loads at one week post-inoculation than macaques inoculated with SHIV\textsubscript{Vif5A}, however the virus was effectively controlled by the macaques by 3 weeks post-inoculation. Similar to our previous study, the majority of the G-to-A mutations were in the context of 5'-TC (minus strand) and not 5'-CC. In addition, we observed that the majority of the G-to-A changes found in the \textit{vif} gene amplified from PBMC DNA occurred during the first weeks of infection. This correlated well with the plasma viral loads, which indicated the viral replication primarily occurred during this time period.

Taken together, our results show that abolishing the SLQYLA domain of Vif may be more detrimental to the virus \textit{in vivo} than the HCCH domain. These data bring up the question, “Why were we able to quantify viral RNA from tissues of macaques inoculated with SHIV\textsubscript{VifHCCH(−)} but not SHIV\textsubscript{Vif5A}?” While the answer is unknown, several possibilities exist. First, the virus may have initially replicated in cells expressing active rhA3
proteins and incorporated one or more rhA3 proteins into the viral progeny. In the next round of replication the genome was mutated but not to the extent that RNA polymerase II could not transcribe viral RNA. Second, there may be a tissue/cellular reservoir for virus replication that do not express the rhA3 proteins that would restrict the crippled SIV\textsubscript{mac}239 Vif protein. We previously showed using immunohistochemistry that rhA3G was not expressed in all cell types in the brain and kidneys (Hill et al., 2006; 2007). Finally, during the initial rounds of replication and potential deamination, G-to-A mutations may have led to compensating mutations that made the viral Vif partially functional. The later scenario is not likely as our sequence analysis did not identify consensus mutations in either the SHIV\textsubscript{Vif5A} or SHIV\textsubscript{VifHCCH(-)}-inoculated macaques.

As previous studies showed that antibody responses were not generated with a single inoculation of \textit{vif}-deleted viruses (Desrosiers et al., 1998; Sparger et al., 2008), we determined if a single inoculation of either virus would result in an antibody response against viral proteins. Comparison of the antibody responses from macaques inoculated with either SHIV\textsubscript{VifHCCH(-)} or SHIV\textsubscript{Vif5A} shows that at 12 weeks post-inoculation macaques had developed antibody responses against the virus but these were undetectable at necropsy. This would argue that the viral RNA detected in tissues at necropsy may not have been translated into viral proteins. This would also suggest that infectious virus was cleared by the macaque and that antigenic stimulation through viral proteins did not occur. However, this study provides evidence that such viruses containing site-directed mutations in \textit{vif} could be used to “prime” the immune system and suggests that such a virus could be useful in a prime-boost strategy. It will be of interest to determine if multiple inoculations will result in stronger immune responses (both humoral and cellular).
and a further reduction of viral loads with increasing inoculations. If successful, it will be of interest to determine if the immune responses are protective against challenge. While the use of live attenuated lentiviruses has led to useful information concerning the role of accessory genes in replication of the virus in vivo, the use of such vaccines has several underlying risks (Koff et al., 2006). First, these vaccines have relied on the deletion/disruption of accessory genes (nef, vpu, vpx/vpr) that enhance replication but are not absolutely required for replication in vivo. For this reason these viruses can select for compensating mutations that ultimately result in the virus becoming pathogenic (Baba et al., 1995; 1999). The second problem is the risk of recombination of the viral genome to yield a pathogenic virus (Kim et al., 2005). A vaccine based on a “crippled Vif” that allows for limited replication resulting in G-to-A substitutions and inactivation of the virus should prevent the scenarios discussed above. It would be of interest to determine if inoculation of macaques with two viruses, one harboring a mutation in vif and another having a mutation in another accessory gene such as nef can recombine in vivo to generate a wild type virus causing CD4⁺ T cell loss and disease prior to the virus accumulating G-to-A mutations.

**Experimental Methods**

**Cells, plasmids, and viruses**

The C8166 and SupT1 lymphocyte cell lines were used for transfections of full-length SHIVs and as indicator cells to measure infectivity and cytopathicity of the viruses
used in this study. Both cell lines were maintained in RPMI-1640, supplemented with 10 mM Heps buffer pH 7.3, 2 mM glutamine, 5 µg per ml gentamicin, 100 units/µg penicillin-streptomycin and 10% fetal bovine serum (R10FBS). Rhesus macaque PBMCs were obtained from uninfected animals and isolated on Ficoll/Hypaque gradients for p27 growth curves. The 293 cell line was maintained in Dulbecco’s minimal essential medium (DMEM) with 10% fetal bovine serum and the antibiotics described above. The derivation of SHIV<sub>KU-2MC4</sub> has been previously described (Joag et al., 1998). The pcDNA3.1(+)-HA-rhesus APOBEC3G and pcDNA3.1(+)-rhesus APOBEC3F-V5 were kindly provided by Nathaniel Landau (New York University School of Medicine, New York, New York).

**Construction of SHIV<sub>VifSA</sub>, SHIV<sub>VifHCCH(-)</sub> and SHIV<sub>VifSTOP</sub>**

For the construction of SHIV<sub>VifHCCH(-)</sub>, the PacI/SphI fragment (nucleotides 5132 to 6452) from the p5'-SHIV4 was subcloned into the pGEM3Zf (+) vector. The histidine to alanine substitution at position 110 of Vif was introduced using oligonucleotides (only sense strand shown) 5'-GCAGACATTTTACTGGCTAGCACTTATTTCC-3'. The cytidine to alanine substitution at position 116 of Vif was introduced using oligonucleotides (only sense strand shown) 5'-GCACTTATTTCCCTGCCTTTACAGCGGGAG-3'. The cysteines at positions 134 and 135 of Vif were substituted for alanines using the oligonucleotides (only sense strand shown) 5'-CAACTGCTGTCTGCCGCCAGGTTCCCG-3'. The histidine to alanine substitution at position 144 of Vif was introduced using oligonucleotides (only sense strand shown) 5'-GGTTCCCGAGAGCTGCTAAGTACCAGGTACC-3'. All substitutions were made using
the Quick-Change Mutagenesis Kit (Stratagene) following the manufacturer's instructions. The resulting PacI/SphI fragment was digested, isolated, and subcloned into full length SHIV\textsubscript{KU-2MC4}. The resulting plasmid was sequenced to ensure that the desired mutations were introduced as expected.

For the construction of SHIV\textsubscript{Vif5A}, the PacI/SphI fragment (nucleotides 5132 to 6452) from full-length SHIV\textsubscript{KU-2MC4} was subcloned into the pGEM3zf (+) vector. The serine and leucine at positions 147 and 148 were substituted for alanines using the oligonucleotides (sense strand shown) 5’-CCAGGTACCAGCCGCACAGTACTTAGCAC-3’. The glutamine and tyrosine at positions 149 and 150 were substituted for alanines using the oligonucleotides (sense strand shown) 5’-CCAGGTACCAGCCGCAGCGGCCTTAGCAGTGAAGTAGTAAGC-3’. The leucine at position 151 was substituted for an alanine using the oligonucleotide (sense strand shown) 5’-GGTACCAGCCGCAGCGGCCGCAGCACTGAAGTAGTAAGCG-3’. All substitutions were made using the Quick-Change Mutagenesis Kit (Stratagene) following the manufacturer’s instructions and virus constructed and prepared as described above.

For the construction of SHIV\textsubscript{VifSTOP}, the tyrosine and leucine at amino acid positions 28 and 29 were engineered to have stop codons using site-directed mutagenesis in order to produce a full-length SHIV that does not express a functional Vif protein. The oligonucleotide used to introduce these mutations was 5’-GCCTCATTAAAATAGTAGAAATATAAAACTAAAG-3’ (only sense strand shown). The substitutions were made using the Quick-Change Mutagenesis Kit (Stratagene) following the manufacturer’s instructions. The resulting PacI/SphI fragment was digested,
isolated, and subcloned into full length SHIV_{KU-2MC4}. The resulting plasmid was sequenced to ensure that the desired mutations were introduced as expected. For all three viruses, the plasmids containing the full-length genomes were transfected into SupT1 cells as previously described (Stephens et al., 2002; Hout et al., 2005). Stocks of SHIV_{VifHCCH(-)}, SHIV_{Vif5A}, and SHIV_{VifSTOP} were prepared, titered on SupT1 cells, and stored at -86°C.

**Analysis of the replication of SHIV_{Vif5A} and SHIV_{VifHCCH(-)} in APOBEC3G/F positive and negative cells**

C8166 (A3G/F positive) and SupT1 (A3G/F negative) cells were inoculated with equivalent levels (25 ng) of parental SHIV_{KU-2MC4}, SHIV_{Vif5A}, SHIV_{VifHCCH(-)}, SHIV_{VifAAQYL}, or SHIV_{VifSTOP} for 4 hours at 37°C. For rhesus PBMC, cells were isolated on Ficoll-Hypaque gradients, stimulated for 48 hours in R10FBS supplemented with concanavalin A (10 μg/ml) and interleukin-2 (IL-2; 50 ng/ml). The cells were washed, and inoculated with virus (25 ng) incubated in R10FBS containing interleukin-2 (50 ng/ml) for 4 hours. At 4 hours, cells were centrifuged, washed 3 times to remove the inoculum and incubated in fresh medium (medium for the rhesus PBMC also contained 50 ng/ml IL-2) at 37°C for up to 15 days. Aliquots of culture supernatants were obtained at 0, 1, 3, 5, 7, 9, 11, 13, and 15 days post-inoculation and the levels of p27 assessed using commercial p27 antigen capture kits (Zeptometrix).

**APOBEC3G/F incorporation assays**

Plasmids containing the genomes viruses (derived from SHIV_{KU-2MC4}) expressing each of the mutant Vif proteins described above were co-transfected into 293 cells using
PEI Transfection reagent (Ex-Gen 500) along with plasmids expressing either HA-rhA3G or V5-rhA3F. At 48 hours, the virus containing supernatants were harvested and clarified by low speed centrifugation. The clarified supernatant was then subjected to ultracentrifugation to pellet the virus (SW41 rotor, 247,000xg, 1 hour). The pellet was resuspended in PBS (pH 7.4) and layered on a 20/60% sucrose step gradient and again subject to ultracentrifugation (SW55Ti, 247,000xg, 1 hour). The virus (at the interface) was harvested, pelletted again by ultracentrifugation described above, and resuspended in 200µl of 1x PBS (pH 7.4). An aliquot was saved to determine the p27 content by antigen capture assay (Zeptometrix). The remaining sample was boiled in sample reducing buffer. Equivalent amounts of p27 were loaded on a 12% SDS-PAGE gel and transferred to PVDF membranes. APOBEC3 proteins were detected by Western blotting using an antibody directed against either the HA tag (HA-probe; Santa Cruz) or V5 tag (Clone V5-10; Sigma). Blots were stripped in 1X stripping buffer (62.5 mM Tris-HCl, pH 6.8 and 2% SDS) and reprobed using a rabbit polyclonal antibody specific for p27. As a control, plasmids expressing either rhA3G or rhA3F were transfected into 293 cells using PEI transfection reagent (ExGen500, Fermentas). At 48 hours post-transfection, cells were lysed in 1x RIPA and the nuclei were removed. Whole protein was precipitated with methanol and resuspended in 2x sample buffer. Normalized to β-actin, the samples were run on a 12%-SDS-PAGE gel, and probed with the antibodies stated above.

**Stability of rhA3G and rhA3F in the presence of SHIV genomes**
We determined if rhA3G or rhA3F were stable in the presence of SHIV\textsubscript{KU-2MC4}, SHIV\textsubscript{VifHCCH(-)}, SHIV\textsubscript{Vif5A}, or SHIV\textsubscript{VifSTOP}. Full-length mutated SHIVs were co-transfected in a 2:1 ratio with either APOBEC3G-HA or APOBEC3F-V5 using polyethylenimine transfection reagent (PEI, Fermentas). Twenty-four hours post-transfection, the supernatant was removed and the cells were harvested and lysed using 1 x RIPA (50 mM Tris-HCl, pH 7.5; 50 mM NaCl; 0.5% deoxycholate; 0.2% SDS; 10 mM EDTA). Following lysis, the nuclei were removed by microcentrifugation at 1400 rpm for 15 minutes. The protein was prepared using methanol, resuspended in 2X sample reducing buffer, and boiled for 5 minutes. Proteins were separated on a 10% SDS-PAGE gel and probed using commercially available rabbit polyclonal HA antibody (HA-probe, Santa Cruz) for rhA3G-HA or mouse monoclonal V5 antibody (Clone V5-10, Sigma) for rhA3F. All samples were normalized to the amount of β-actin protein using a mouse monoclonal antibody (Novous Biologicals) specific for β-actin.

**Hypermutation assays in the presence of rhA3G or rhA3F**

Hypermutation of various SHIVs in the presence of rhesus APOBEC3G/F 293 cells were transfected with SHIV\textsubscript{KU-2MC4}, SHIV\textsubscript{VifHCCH(-)}, SHIV\textsubscript{Vif5A}, or SHIV\textsubscript{VifSTOP} in the presence of either rhA3G or rhA3F using polyethylenimine transfected reagent (PEI, Fermentas). Twenty-four hours post-transfection, cells were washed and fresh DMEM was added. After forty-eight hours, the supernatant containing virus was subjected to low speed centrifugation. The resulting supernatant was DNase-I—treated (Fermentas) at 37°C for 30 minutes to eliminate any trace of plasmid carry-over from the initial
transfection. The DNase-I-treated supernatant was titrated on TZM-bl cells to both assess infectivity and hypermutation. Twenty-four hours post-infection, total DNA cellular DNA was harvested and extracted using the DNeasy kit and the manufacturer's instructions (Qiagen). The DNA was used in a nested DNA polymerase chain reaction to amplify a 300 base pair fragment of nef. The PCR reaction was carried out using rTaq, the manufacturer's instructions (Takara), and the oligonucleotides listed below. 1 µl of the first PCR product was added to a nested reaction. The PCR reactions were performed using an Applied Systems 2720 Thermal Cycler with the following thermal profile: 95°C for 2 minutes, 1 cycle; 95°C for 30 seconds, 48°C for 30 seconds, 65°C for 2 minutes, 35 cycles; 65°C for 7 minutes. The PCR products were separated by electrophoresis, isolated, purified, sequenced and sub-cloned into pGEM-TEasy (Promega) as described below. Fifteen independent clones were sequenced and assessed for each mutant SHIV as described above.

Restriction of Vif Mutants on rhA3G or rhA3F

We determined the effect of the Vif mutants on the suppression of rhA3G and rhA3F. Full-length mutated SHIVs were co-transfected with either rhA3G or rhA3F using PEI transfection reagent (ExGen500, Fermentas). Forty-eight hours post-transfection, the supernatant was harvested and purified by low speed centrifugation. Equivalent amounts of p27 were serially diluted using 10-fold dilutions from 10^4 to 10^6 and used to inoculate TZM-bl cells. Forty-eight hours later, the media was removed, cells washed with 1X PBS and the monolayer fixed using 1% formaldehyde-0.2% glutaraldehyde in
phosphate-buffered saline (1 x PBS). The cells were washed and incubated in a solution for 2 hours at 37C in 4 mM potassium ferrocyanide, 4 mM potassium ferricyanide, 4 mM magnesium chloride, and 0.4 mg X-gal per ml. The reaction was stopped and the number of TCID$_{50}$ were calculated (Derdeyn et al., 2000; Wei et al., 2002).

**Macaques analyzed in this study**

Three rhesus macaques (*Macaca mulatta*: CX54, ER65, and I92) were intravenously inoculated with 1 ml of undiluted culture supernatant from SupT1 grown virus stocks (containing $10^4$ TCID$_{50}$ per ml). Three additional rhesus macaques (AS34, AS51, and AV18) were inoculated with $10^4$ TCID$_{50}$ SHIV$_{VifHCCH(-)}$ (titered in SupT1 cells). The animals were housed in the AAALAC-approved animal facility at the University of Kansas Medical Center. All aspects of the animal studies were performed according to the institutional guidelines for animal care and use at the University of Kansas Medical Center. Heparinized blood was collected weekly for the first 4 weeks, then at 2-week intervals for the next month and at monthly intervals thereafter.

**Assays for circulating CD4$^+$ T cells**

Changes in the levels of CD4$^+$ T cells after viral inoculations were monitored sequentially by flow cytometric analysis (BD Biosciences). T cell subsets were labeled with a commercially available anti-CD3/CD4/CD8 mixture. T cell subsets from a normal uninfected macaque were always performed at the same time to serve as a control for the flow cytometry analysis.
Processing of tissue samples at necropsy

At the time of euthanasia (26 or 28 weeks), all macaques in this study were anesthetized by administration of 10 mg/kg ketamine (IM) followed by intravenous administration of sodium phenobarbital at 20-30 mg/kg. At the time of necropsy, all macaques were in a healthy condition. A laparotomy was performed. The animal was exsanguinated by aortic canulation and perfused with one liter of cold Ringer’s saline. Lymphoid and non-lymphoid tissues (heart, kidneys, liver, lungs, mesenteric, inguinal and axillary lymph nodes, pancreas, salivary gland, small intestine, spleen, thymus, and tonsils) were obtained and aliquots of tissue snap frozen for DNA and RNA assays. Aliquots of lymphoid tissues were immersed in HBSS to quantify levels of infectious virus in tissues using infectious centers assays with SupT1 cells.

Sequence analysis of the vif, nef, and vpu genes

To determine the stability of the mutations that were introduced into vif and assess whether these macaques acquired G-to-A mutations, the vif, nef, and vpu genes were amplified from either PBMCs at different time points during infection (vif) or from several tissue DNA samples taken at necropsy (nef and vpu) that were positive for viral RNA. One hundred nanograms of extracted genomic DNA was used in a nested DNA polymerase chain reaction (Takara, Madison, WI) following the manufacturer’s instructions. The oligonucleotides employed during the first round to amplify vif were 5’-GGCTAAAATTATCAAGATTATGGAGG-3’ (sense) and 5’-
GGTGACATCCCTGTTCATCATGCC-3’ (antisense), which corresponds to bases 5326-5352 and 5984-6008, respectively. The nested oligonucleotides were 5’-GGAGGAGAAAGAGGTGGATAGCAGTTCCC-3’ (sense) and 5’-CCAGTATTCCCAAGACCTTTGCC-3’ (antisense), which corresponds to bases 5348-5378 and 5963-5985, respectively. The oligonucleotides used during the first round to amplify the nef gene were 5’-GGTGGAGCTATTTCCCATGAGG-3’ (sense) and 5’-GTCTTCTTGGACTAATAATCC-3’ (antisense), which corresponds to bases 9445-9465 and 9832-9856, respectively. The nested oligonucleotides were 5’-CCATGAGGCGGTCCAGGCAGTCTAGAG-3’ (sense) and 5’-CCTCCCAGTCCCCCCTTTTC-3’ (antisense), which corresponds to bases 9458-9484 and 9814-9833, respectively. The oligonucleotides used during the first round to amplify the vpu gene were 5’-CCTAGACTAGAGCCCTGGAAGCATCC-3’ (sense) and 5’-GTACCTCTGTATCATATGCTTTAGCAT-3’ (antisense), which corresponds to bases 6486-6511 and 7034-7061, respectively. The nested oligonucleotides used were 5’-TTAGGCATCTCCTATGGCAGGAAGAAG-3’ (sense) and 5’-CACAAAATAGAGTCCCTGGTTGCTTC-3’ (antisense), which corresponds to bases 6597-6623 and 7001-7027, respectively.

For sequence analysis, the PCR products from three separate PCR reactions were pooled and separated by electrophoresis in a 1.5% agarose gel, isolated, and each PCR reaction directly sequenced. Cycle sequencing reactions using the BigDye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq DNA polymerase, FS (PE Applied Biosystems, Foster City, CA) sequence detection was conducted with an
Applied Biosystems 377 Prism XL automated DNA sequencer and visualized using the ABI Editview program. Sequences were compared to the sequence of SHIV\textsubscript{KU-2MC4}. A total of 615 nucleotides were analyzed from \textit{vif}, 300 nucleotides analyzed from \textit{nef}, and 184 nucleotides from \textit{vpu} using the SE Central Software package. In order to isolate and analyze single sequences, bulk PCR reactions were subcloned into the pGEM-TEasy (Promega) cloning vector. Fifteen clones were selected, sequenced as described above, and analyzed to assess the number of G-to-A substitutions that occurred.

**Plasma virus loads**

Plasma viral RNA loads were determined on RNA extracted from 1 ml of EDTA-treated plasma. Virus was pelleted using ultracentrifugation (Beckman SW55Ti, 250,000xg, 2 hours) and RNA extracted using the Qiagen viral RNA kit (Qiagen). RNA samples were analyzed by real-time RT-PCR using \textit{gag} primers and a 5'FAM and 3'TAMRA labeled Taqman probe that is homologous to the SIV \textit{gag} gene as previously described (Hofmann-Lehmann \textit{et al.}, 2000). Standard curves were prepared using a series of six ten-fold dilutions of viral SIV \textit{gag} RNA of known concentration. The sensitivity of the assay was 100 RNA equivalents per ml. Samples were analyzed in triplicate and the number of RNA equivalents per ml of plasma were calculated. Viral RNA was also quantified from visceral tissues using this primer/probe reaction. In order to quantify the amount of viral RNA, GAPDH was used as a control and data is presented as copies per $10^6$ GAPDH molecules (Marcario \textit{et al.}, 2008).
**Immunoprecipitation assays**

To determine if the macaques developed antibodies to SHIV proteins following inoculation, the plasma at 12 weeks and at necropsy was used in immunoprecipitation assays. C8166 cells were inoculated with approximately $10^4$ TCID$_{50}$ of SHIV$_{KU-2MC4}$ for 5 days. The cells were then incubated in methionine/cysteine-free media for 2 hours and radiolabeled labeled with 500 µCi of $^{35}$S-methionine/cysteine for 15 hours. The cells were lysed in 1ml of 1X RIPA buffer and nuclei were removed by centrifugation. The cell lysates were incubated overnight at 4°C with 10 µl of plasma from each macaque and protein A Sepharose beads. The immunoprecipitates bound to the beads were washed three times in 1X RIPA, resuspended in 75 µl of 2X sample reducing buffer, and boiled for 5 minutes. Proteins were separated on a 10% SDS-PAGE gel and visualized by autoradiography. Controls included pooled prebleed plasma from macaques (negative control) and plasma from macaques that had been previously inoculated with a non-pathogenic SHIV (SHIV$_{TM}$, positive control).
VIII. Chapter 4: Differential virus restriction patterns of rhesus macaque and human APOBEC3A: Implications for lentivirus evolution

Abstract

The human apolipoprotein B mRNA editing enzyme catalytic peptide-like 3 (APOBEC3; A3) family of proteins (A3A-H) are known to restrict various retroviruses and retroelements, but the full complement of rhesus macaque A3 proteins remains unclear. We report the isolation and characterization of the hA3A homologue from rhesus macaques (rhA3A) and show that the rhesus macaque and human A3 genes are orthologous. RhA3A is expressed at high levels in activated CD4+ T cells, is widely expressed in macaque tissues, and is degraded in the presence of the human immunodeficiency virus (HIV-1) and simian-human immunodeficiency virus (SHIV) genomes. Our results indicate that rhA3A is a potent inhibitor of SHIVΔvif and to a lesser extent HIV-1Δvif. Unlike hA3A, rhA3A did not inhibit adeno-associated virus 2 (AAV-2) replication and L1 retrotransposition. These data suggest an evolutionary switch in primate A3A virus specificity and provide the first evidence that a primate A3A can inhibit lentivirus replication.

Introduction

In the last decade several host restriction factors have been discovered that can restrict the replication of HIV-1. One of these restriction factors, Apolipoprotein B
mRNA-editing catalytic polypeptide-like 3 (APOBEC3; A3), has been shown to inhibit a wide range of retroviruses and other viruses such as the parvoviruses, papillomaviruses, and hepadnaviruses (Abe et al., 2009; Baumert et al., 2007; Bonvin et al., 2006, 2007; Chen et al., 2006; Henry et al., 2009; Jarmuz et al., 2002; Köck et al., 2008; Mahieux et al., 2005; Narvaiza et al., 2009; Noguchi et al., 2007; Paprotka et al., 2010; Sheehy et al., 2002; Strebel et al., 2005; Turelli et al., 2004; Vartanian et al., 2008; Zhang et al., 2008). A3 proteins comprise a family of seven cytidine deaminases (A3A, A3B, A3C, A3D, A3F, A3G, and A3H) that catalyze the deamination of cytidine to uracil on single-stranded DNA or RNA through a highly conserved Zn^{2+}-binding motif (C/H-X-E-X_{23-28}-P-C-X_{2-4}-C) (Chiu et al., 2009; Goila-Gaur and Strebel., 2008; Huthoff et al., 2005; Yu et al., 2004). These proteins also contain a key glutamate required for proton shuttling during catalysis and two aromatic amino acids residues required for RNA substrate binding (Jarmuz et al., 2002).

The A3 proteins can be divided into those with a single (A3A, A3C, A3H) or double (A3B, A3D, A3F, and A3G) cytidine deaminase domains (CDA). The double CDA proteins A3D, A3F, and A3G have been shown to restrict the replication of HIV-1 isolates that do not express a functional Vif protein while A3B could inhibit HIV-1 independent of Vif (Dang et al., 2006; Doehle et al., 2005; Kao et al., 2003; Liddament et al., 2004; Sheehy et al., 2002; Smith et al., 2010; Wiegand et al., 2004; Yu et al., 2004; Zheng et al., 2004). The Vif protein is known to bind to select A3 proteins and shunt the A3 proteins to the proteasome for degradation (Liu et al., 2004; Marin et al., 2003; Mehle et al., 2004; Sheehy et al., 2003; Yu et al., 2003). Mutagenesis studies have shown that the N-terminal CDA lacks catalytic activity and is responsible for
binding the RNA while the C-terminal CDA mediates sequence-specific deamination of single-stranded DNA (Yu et al., 2004). In addition to deaminase dependent functions, hA3G and hA3F inhibit virus replication by deaminase-independent mechanisms (Holmes et al., 2007a,b). Human A3G (hA3G) and hA3F can also inhibit DNA synthesis, which may involve interference of t-RNA Lys3 primer annealing, initiation and elongation of DNA synthesis and minus or plus strand DNA transfers (Bishop et al., 2008; Guo et al., 2006, 2007; Holmes et al., 2007a; Iwatani et al., 2007; Li et al., 2007; Luo et al., 2007; Newman et al., 2005; Ooms et al., 2010). Of the single CDA A3 proteins, hA3A has no activity against HIV-1 while hA3C can be incorporated into virions but only has weak activity against HIV-1 and foamy virus (Langlois et al., 2005; Perkovic et al., 2009). Human A3H is now known to have seven haplotypes, several of which can be stably expressed (Harai et al., 2009; Ooms et al., 2010; Wang et al., 2011). However, attempts to detect hA3H in isolated peripheral blood mononuclear cells (PBMC) has been unsuccessful (Li et al., 2010).

Human A3A is a nucleocytoplasmic protein that is expressed mainly in primary monocytes and keratinocytes (Koning et al., 2009; Madsen et al., 1999; Peng et al., 2007). Sequences consistent with hA3A-mediated editing have been observed in the genome of papilloma virus, which is known to infect both cutaneous and mucosal keratinocytes (Vartanian et al., 2008). Previous studies shown that hA3A lacks antiviral activity against HIV-1, but inhibits the replication of adeno-associated virus 2 (AAV-2), intracisternal A particles (IAP), and long interspersed nuclear element 1 (L1) (Bogerd et al., 2006a,b; Chen et al., 2006; Muckenfuss et al., 2006). The expression of HA3A is
induced by foreign DNA detection and interferon (IFN) in phagocytes, which leads to the degradation and clearance of foreign DNA (Stenglein et al., 2010).

Although hA3A contains one CDA, the deaminase activity of hA3A is not required for its inhibitory role against AAV or retrotransposons, suggesting that a deaminase-independent mechanism of inhibition is occurring (Narvaiza et al., 2009). Human A3A also contains significant homology to the N-terminal half of hA3G, does not associate with the viral nucleocapsid complex (NPC) of HIV-1 but is readily incorporated into virions (Goila-Gaur et al., 2007).

Simian immunodeficiency (SIV) and simian-human immunodeficiency (SHIV) viruses have been invaluable models for studying various aspects of HIV-1 pathogenesis. Sequencing of the rhesus macaque genome has led to the identification of rhesus homologues of A3B, A3C, A3D, A3F, A3G and A3H. Here we describe the identification of an A3A homologue in rhesus macaques (rhA3A) and present evidence that the human and macaque Apobec3 genes are orthologous. We show that the rhA3A protein is expressed well in activated macaque CD4+ T cells and other rhesus macaque tissues, is incorporated into virions, is degraded by both HIV-1 and SHIV in a Vif dependent manner, and restricts the replication of SHIV (expressing the SIVmac239 Vif) and HIV-1.

Results

Molecular cloning of rhesus macaque A3A
Humans encode seven APOBEC3 (hA3) genes that are tandemly arrayed on chromosome 22, but there is uncertainty on whether a similar array of rhA3 genes is expressed in the syntenic region on rhesus macaque chromosome 10. In particular, no rhA3A gene has been cloned from rhesus macaques suggesting that this gene may have been deleted in the genome of this primate species (OhAinle et al., 2008; Virgen et al., 2007). However, a recent study showed that A3A is encoded in the genome of several monkey species, including African green and owl monkeys (Bulliard et al., 2011). The evolutionary conservation of A3A therefore suggested that the rhA3A gene was missed in previous cloning attempts. To determine whether rhA3A is expressed, we took advantage of the sequencing of the rhesus macaque genome (Rhesus Macaque Genome Sequencing and Analysis Consortium, 2007). BLAST analysis of the hA3A sequence revealed that the first exon has a counterpart on the rhesus macaque chromosome 10, thus providing a potential forward primer for PCR amplification. However, counterparts of other hA3A exons were not detected in the rhesus macaque genome immediately downstream of this putative exon 1 sequence. Thus, the last exon of hA3A was used as a reverse primer.

Human A3A is not expressed in resting peripheral blood mononuclear cells (PBMC), but is substantially induced by cellular activation and cytokines (Koning et al., 2009; Refsland et al., 2010). We therefore utilized cDNA from Concanavilin A-activated rhesus macaque PBMCs in an attempt to detect rhA3A. A dominant band of the correct size was amplified in the first attempt by PCR. BLAST analysis of the cloned amplicon revealed 89% nucleotide identity with hA3A. Alignment of the amino acid sequence with recently identified A3A sequences showed the highest identity (90%) with A3A from
Figure 41. Sequence analysis of the rhA3A proteins. Panel A. Protein sequence alignment of primate A3A proteins. Dashes correspond to indels, while residues highlighted in pink or orange are conserved or matched the consensus, respectively. Some indels were associated with specific primate taxons. This includes homiind-specific deletion at amino acid positions 27 to 29, and a deletion in residues 108-109 in New World monkeys. Panel B. A3A phylogeny. The amino acid sequence alignment was subjected to phylogenetic analysis using the neighbor joining method, excluding sequence gaps, and rooted using the C-terminal half of rhesus macaque A3G as an outgroup. Numbers correspond to the percentage of clustering following 1000 subreplicates. RhA3A clustered most significantly with A3A from another Old World monkey species, African Green monkey (AGM).
African green monkeys, which, together with Asian macaques, are considered as “Old World” monkeys (Figure 41). This alignment also revealed primate lineage-specific insertions/deletions (indels). In particular, we observed deletions at positions 27 to 29 in the hominid A3A genes. To further explore the clustering of A3A sequences with specific primate lineages, we constructed a phylogenetic tree that excluded the contribution of indels. As shown in Figure 41, A3A sequences clustered accordingly with their corresponding primate taxons. These analyses demonstrate that the cloned rhA3A gene is a *bona fide* member of the A3A gene family.

**Evidence that seven A3 genes are tandemly arrayed on chromosome 10 of rhesus macaques**

While the sequence of the rhesus macaque genome has been “completed,” it still remains unclear why this did not reveal the *rhA3A* gene. We attempted to determine the position of the cloned *rhA3A* sequence in relation with other *rhA3* genes. At the time that the analyses were performed, only the nucleotide sequences for rhesus macaque *rhA3C, rhA3F, rhA3G* and *rhA3H* had been deposited into GenBank. BLAST analysis of these cDNA sequences against the rhesus macaque genome revealed their distinct positions as shown in Figure 42. The position of rhesus macaque *rhA3B* and *rhA3DE* remains unknown, although their amino acid sequences were published (Virgen et al., 2007). We first attempted to determine the relative positions of *rhA3B* and *rhA3D* in chromosome 10 based solely on the genome sequence. Using the geneid program (http://genome.crg.es/geneid.html), genome sequences between *rhA3C* and *rhA3F*
predicted *rhA3D* (Figure 43), while genomic and previously unlocalized sequences upstream of *rhA3C* predicted *rhA3B* (Figure 44).

The relative placement of *rhA3B* and *rhA3D* in the rhesus macaque genome suggested a similar A3 gene organization as found in humans, with the notable uncertainty of *rhA3A*. We therefore focused on genomic sequences immediately upstream of *rhA3B*. BLAST analysis of the cloned *rhA3A* sequence revealed an identical sequence corresponding to *rhA3A* exon 1 in the genome. This putative *rhA3A* exon 1 was immediately followed by 560 bp of CT-rich repeat sequences and 16.8 kb of unresolved sequence. It is possible that these repeat sequences interfered with accurate sequencing of this region. On the other hand, exons 2 to 4 of *rhA3A* mapped to an unlocalized chromosome 10 genomic scaffold sequence with GenBank Accession NW_001096321.1. The 3,764 bp NW_001096321.1 sequence could easily be accommodated in the 16.8 kb genomic gap. Notably, the NW_001096321.1 sequence contained 838 bp of intron sequence upstream of exon 2. We therefore investigated whether the NW_001096321.1 links directly with the genome by amplifying rhesus macaque genomic DNA with primers designed in exon 1 and 2 of *rhA3A* (Figure 42). Direct sequencing of the 1.8-kb amplicon revealed identity with the flanking sequences following the putative exon 1 in the genome and prior to exon 2 in NW_001096321.1 (Figure 42). These results provide strong evidence that *rhA3A* is likely encoded in the rhesus macaque genome immediately upstream of *rhA3B*. Overall, these analyses suggest that the seven *rhA3* members are tandemly arrayed in the rhesus macaque genome, similar to humans.
Figure 42. Sequence characterization of the rhesus macaque A3 locus. Panel A. Genome organization of rhesus macaque A3 genes. *(Top)* Published cDNA sequences for rhA3C, rhA3F, rhA3G, and rhA3H were subjected to BLAST analysis against the rhesus macaque genome. Matching exons were plotted against assigned genome coordinates, shown here in 20 kb intervals. *RhA3B* and rhA3D were predicted by subjecting the genome sequences upstream from rhA3C or flanked by rhA3C and rhA3F to the geneid prediction program ([http://genome.crg.es/geneid.html](http://genome.crg.es/geneid.html)), respectively. RhA3A amplified from activated PBMC was subjected to BLAST analysis. RhA3A exon 1 mapped in the genome directly upstream of A3B, but this was followed by a 16.8 kb of unresolved sequence (dotted lines). *(Bottom)* A3A exons 2 to 4 matched sequences from an unlocalized chromosome 10 genomic scaffold sequence. Primers based on A3A exons 1 and 2 amplified a 1.8 kb product that matched flanking sequences after exon 1 in the genome and before exon 2 in the unlocalized genomic scaffold (gray lines). This provides strong evidence for the current placement of rhA3A in the rhesus macaque genome. Panel B. Phylogeny of rhesus macaque and human A3 proteins. Amino acid sequences from rhesus macaque and human A3 genes were aligned, with the double-CDA members split into an N- and C-terminal half. For example, A3B-C corresponds to the C-terminal half of A3B. RhA3A clustered with the C-terminal half of A3B and A3G, forming the Z1 clade. Each rhesus macaque A3 gene clustered significantly with the corresponding human A3 gene, suggesting that rhesus macaque and human A3 proteins are orthologous.
Figure 43. Nucleotide sequence alignment of hAu3D and genome-predicted rhA3D. The rhesus macaque genome sequence flanking rhA3C and rhA3F could potentially encode rhA3D. To investigate this possibility, we subjected this genomic sequence to a gene prediction program, (http://genome.crg.es/geneid.html). This analysis revealed a putative rhA3D gene that exhibited 86% nucleotide identity with hA3D. However, multiple attempts to clone the full-length gene in activated rhesus macaque PBMCs failed, likely due to the fact that the reverse primer cross-reacted significantly with the more highly expressed rhA3C gene and that rhA3D is very poorly expressed in these cells. In contrast, the 5’ half of this gene was cloned (unpublished data).
Figure 44. Nucleotide sequence alignment of hA3B and genome-predicted rhA3B. Upstream of rhA3C are genome sequences that matched some exons of hA3B. Closer inspection revealed unresolved sequences in regions that could putatively encode exons 2, 7, and 8. A substantial portion of these gaps was resolved with the inclusion of “unlocalized genomic scaffold sequences” from chromosome 10. In particular, chromosome 10 sequences designated with GenBank Accession numbers AANU01162607.1 and AANU01065938.1 encoded putative exons 2 and 7 of A3B and formed contiguous bases to bridge a portion of the unresolved regions. A 2.7-kb gap that likely encodes the last 4 amino acids of A3B still remains unresolved in the macaque genome. Using the geneid program, the rhA3B nucleotide sequence was predicted, and this sequence exhibited 85% identity with hA3B. These analyses provided strong evidence for the current position of rhA3B upstream of rhA3C. However, our multiple attempts to clone rhA3B from activated rhesus macaque PBMC mRNA were unsuccessful.

The rhesus macaque and human A3 genes are orthologous
null
Rodents encode a single A3 gene, while humans encode seven. Other mammals, such as dogs, pigs, horses and sheep encode a variety of A3 genes that reveal substantial duplication, unequal cross-over and divergent evolution of individual members and are therefore considered paralogues (Conticello et al., 2005; LaRue et al., 2006). This pattern of “evolutionary shuffling” could be roughly inferred in phylogenetic trees, where A3 segments containing the canonical deoxycytidine deaminase motif can be classified into Z1, Z2 and Z3 clades (LaRue et al., 2006). To extend this classification to rhA3 genes, a phylogenetic tree comparing the human, rodent and rhesus macaque A3 amino acid sequences was constructed (Figure 42B). The rhA3 genes consistently clustered with the hA3 counterparts (Figure 42B). Thus, each of the seven rhesus macaque and human A3 genes appear to originate from a common ancestor, indicating that the genes are orthologous. In other words, we found no evidence for significant “shuffling events” in the A3 locus after the evolutionary split between monkeys and humans.

**Expression of rhA3 genes in rhesus macaque tissues and activated CD4\(^+\) T cells**

We next determined rhA3A mRNA expression in various perfused rhesus macaque organs and tissues. Specific rhA3A primers were used to determine expression in 14 visceral organs and 12 regions of the central nervous system (CNS) by RT-PCR. GAPDH was used to assess the integrity of RNA in the samples (data not shown). Our results indicate that rhA3A was not only widely expressed in the visceral organs but also in different regions of the CNS (Figure 45A). We verified the specificity
of the rhA3A amplification by directly sequencing the PCR amplicons (data not shown). Extensive detection of rhA3A in various lymphoid organs raised the possibility that rhA3A may also be expressed in CD4+ T cells, a primary cellular target of SHIV and SIV infection in vivo.

To determine the relative expression of rhA3A against other rhA3 members, we developed Taqman-based real-time PCR assays to specifically quantify each of the seven rhA3 genes (Table VIII). To evaluate specificity, we first performed an end-point PCR and determined by agarose gel electrophoresis whether the primers could amplify other rhA3 cDNAs. If some cross-reactivity with the primers was observed, we subjected 10^8 rhA3 cDNA copies in a real-time PCR reaction that also includes the probe. The results of these analyses are presented in Table VIII. Most of the assays developed did not cross-react with other A3 family members. In those instances where some cross-reactivity was observed, detection of similar A3 members were least 10^4
Figure 45. Expression of rhA3A in rhesus macaque tissues and CD4\(^+\) T cells.

Panel A. RhA3A expression profile in visceral organs and the CNS. Multiple tissues were obtained from an SIV negative rhesus macaque. RNA from these tissues was analyzed for the presence of rhA3A mRNA by RT-PCR. Amplicons were run on 1.5% agarose gels and visualized by staining with ethidium bromide. Amplicons were directly sequenced and found to correspond to rhA3A. All samples were positive for GAPDH RNA to control for integrity of RNA in the samples (data not shown). Upper gel, rhA3A mRNA from 14 visceral organs. Lower gel, rhA3A mRNA from 12 regions of the central nervous system (CNS). The size markers (in base pairs) are shown to the left.

Panel B. Expression of A3 genes in rhesus macaque CD4\(^+\) T cells. Upper, Activation of rhesus CD4\(^+\) T cells. Rhesus macaque PBMCs were activated with 2.5 µg/ml SEB overnight, and subjected to FACS analysis. FACS histograms of CD25 expression in CD3\(^+\)CD4\(^+\) cells before (resting, blue) and after SEB activation (activated; red) are shown. CD4\(^+\) T cells were magnetically purified by negative selection, and RNA was extracted for quantitative RT-PCR analyses. Lower, rhA3 mRNA levels in activate rhesus macaque CD4\(^+\) T cells, A3 cDNA was subjected to real-time quantitative RT-PCR using primers specific for individual rhA3 genes (Table VIII) Dashed lines correspond to the assay limit of detection (5 input copies/reaction), while error bars correspond to standard deviation from triplicate determinations.
B.

**Gated: CD4⁺ T cells**

- **DONOR 1**: 24.7%
- **DONOR 2**: 17.0%
- **DONOR 3**: 61.5%

**Legend:**
- Blue line: resting
- Red line: activated

**RhA3 mRNA**

(activated CD4⁺ T cells)

<table>
<thead>
<tr>
<th>DONOR</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A3A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3D</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3E</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3F</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3H</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Apopec-3/GAPDH copies)
Table VIII. Rhesus macaque A3 real-time PCR primers and probes.
<table>
<thead>
<tr>
<th>Apobec3</th>
<th>Primers[a]</th>
<th>Primer size (nt)</th>
<th>PCR product (nt)</th>
<th>T&lt;sub&gt;a&lt;/sub&gt; (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RhA3A</td>
<td>Forward: CAAGGCCTAAGAAAGTGTCCTCTTCT</td>
<td>20</td>
<td>201</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Reverse: ATGGCCAGCTCTCAGTGGTTC</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probe: [6-FAM][TGCAGGTTCTCTTCTTGGCAGT][TAMRA-6-FAM]</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RhA3B</td>
<td>Forward: TGCCGTAGAGATGCGACAG</td>
<td>20</td>
<td>224</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Reverse: TTGAAGTGGACTGCTGCTG</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probe: [6-FAM][CCGCAATGCTAAAGAGGATTTG][TAMRA-6-FAM]</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RhA3C</td>
<td>Forward: AGGTGCTCCTCCTCCTGTT</td>
<td>20</td>
<td>197</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Reverse: CCGTCTGGTATCTCCTCA</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probe: [6-FAM][GCACCAAAATGTGATGCTC][TAMRA-6-FAM]</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RhA3D</td>
<td>Forward: GTCTCCGAAACCCAGTGTGA</td>
<td>20</td>
<td>245</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Reverse: AGGCCTCGCTGTAATCCTG</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probe: [6-FAM][CTCTGCTCTCGCCTCAGACAA][TAMRA-6-FAM]</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RhA3F</td>
<td>Forward: GAAGGTGGGCGATTCTGA</td>
<td>19</td>
<td>264</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Reverse: TTTCGCTGAACTCCCTGCTG</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probe: [6-FAM][ACTGGAACAGATTATCGCA][TAMRA-6-FAM]</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RhA3G</td>
<td>Forward: GCCATTAGGAGCCTCCTGGAACA</td>
<td>20</td>
<td>230</td>
<td>60.5</td>
</tr>
<tr>
<td></td>
<td>Reverse: GAGCTCCTGGCTAGAGGAAG</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probe: [6-FAM][AAGCTGCTGGGGCTGACAG][TAMRA-6-FAM]</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RhA3H</td>
<td>Forward: AAGGACCAGCAGAATTTG</td>
<td>20</td>
<td>239</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Reverse: ACCTGGAATCCACACAGAAG</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probe: [6-FAM][AATATCGAGGGGCTGCT][TAMRA-6-FAM]</td>
<td>20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[a] PCR conditions were as follows: 95 °C for 10 min followed by 40 cycles of 95 °C for 15 s and the corresponding annealing temperature (T<sub>a</sub>) for 45 s.
Table VIII. Specificity evaluation of rhA3 primers and probes
<table>
<thead>
<tr>
<th>Primer-probe set</th>
<th>Template cross-reactivity (input: $5 \times 10^8$ copies)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Rh\bar{A}3A$</td>
</tr>
<tr>
<td>$Rh\bar{A}3A$</td>
<td>NT</td>
</tr>
<tr>
<td>$Rh\bar{A}3B$</td>
<td>−</td>
</tr>
<tr>
<td>$Rh\bar{A}3C$</td>
<td>−</td>
</tr>
<tr>
<td>$Rh\bar{A}3D$</td>
<td>−</td>
</tr>
<tr>
<td>$Rh\bar{A}3F$</td>
<td>$6.5 \times 10^3$</td>
</tr>
<tr>
<td>$Rh\bar{A}3G$</td>
<td>−</td>
</tr>
<tr>
<td>$Rh\bar{A}3H$</td>
<td>−</td>
</tr>
</tbody>
</table>

$^a$ $Rh\bar{A}3B$ (340nt) and $rh\bar{A}3D$ (370nt) templates were synthetic oligos based on predicted sequences. −, no cross-reactivity observed based on endpoint-PCR and gel electrophoresis; NT: not tested.
fold lower in magnitude than what would be expected for the gene of interest. Thus, the real-time PCR methods could be used to distinguish the relative levels of the individual A3 members.

Rhesus macaque PBMCs from three uninfected donors were stimulated with Staphylococcal Enterotoxin B (SEB), previously determined to be a potent activator of rhesus macaque T cells (Kakimoto et al., 1999). However, SEB-stimulation still did not result in consistent activation, since the levels of CD25 expression in CD4$^+$ T cells varied between the 3 donors (Figure 45B). Total RNA was extracted from magnetically purified CD4$^+$ T cells from these activated PBMCs, and subjected to real-time RT-PCR. Consistent with our failure to clone \textit{rhA3B} from activated PBMCs, \textit{rhA3B} mRNA was not detected in any donor (Figure 45A). With one exception (\textit{rhA3A} donor 3), the expression of individual \textit{rhA3} genes between the donors did not vary more than 10-fold, and did not appear to correlate with CD25$^+$ expression levels. Since \textit{rhA3G} is well accepted as a biologically relevant restriction factor against SIV (Schmitt et al., 2009; 2010; Schrofelbauer 2006; Sui et al., 2010), we compared the expression of other \textit{rhA3} genes relative to \textit{rhA3G}. Setting the average \textit{rhA3G} level from the 3 donors as 1, only \textit{rhA3A} (excluding donor 3), \textit{rhA3C} and \textit{rhA3H} exhibited mRNA expression levels that are similar in magnitude (1.0-1.5 fold) to that of \textit{rhA3G}. In contrast, \textit{rhA3F} and \textit{rhA3D} were on average expressed at 2.4- and 13.6-fold lower than \textit{rhA3G}, respectively. Thus, \textit{rhA3A}, \textit{rhA3C} and \textit{rhA3H} mRNAs are expressed to similar levels as \textit{rhA3G} in SEB-activated rhesus macaque CD4$^+$ T cells.
Nucleocytoplasmic localization of HA-tagged rhA3A

Previous studies have shown that the cytoplasmic localization of hA3G and the nucelocytoplasmic localization of hA3A could partially explain the differences in the range of retroviruses or retrotransposons that were inhibited by these proteins (Bogerd et al., 2006a,b; Chen et al., 2006; Giola-Gaur et al., 2007; Muckenfuss et al., 2006). To determine whether rhA3A exhibited a similar subcellular localization, we generated a molecular clone of rhA3A with an HA-tag at the amino terminal end for detection purposes. A plasmid expressing either rhA3G or rhA3A (Figure 46A) was transfected into 293 cells and the expression examined by Western blot analysis using an antibody directed against the HA tag. As shown in Figure 46A, 293 cells transfected with the vector expressing rhA3G expressed a protein with an Mr of 45kD while cells transfected with the vector expressing rhA3A expressed a protein with an Mr of 24kD, which reflects the size of the two different proteins (363 amino acids for rhA3G versus 202 amino acids for rhA3A). Using confocal microscopy, we found that rhA3G was observed exclusively in the cytoplasm (Figures 46B, C, D) and that rhA3A was detected in both the cytoplasm and nucleus as it co-localized with a nuclear marker (Figures 46B, E, F, G). Similar to other studies, hA3A was observed rhA3A in both the nucleus and cytoplasm (Figures 46G, H, I). Thus any differences in the retroviral restriction properties of hA3A and rhA3A could not be explained by the subcellular localization.
Figure 46. Expression of an HA-tagged rhA3A. Panel A. 293 cells were transfected with 1.5 µg of HA-tagged rhA3A, rhA3G, or the empty expression vectors. At 48 h, cells were lysed and proteins separated by SDS-PAGE and analyzed by Western blot using an antibody directed against the HA-tag. Sizes of the molecular standards are to the right. RhA3A migrates at about half the molecular weight than rhA3G, consistent with their amino acid sequences. Panel B-J. Subcellular localization of rhA3G, rhA3A, and hA3A. 293 cells were transfected with HA-tagged vectors expressing rhA3A, rhA3G, or hA3A and a vector expressing an eGFP tagged nuclear marker (Clontech). At 48 h, cells were fixed and stained with anti-HA for 1 h, washed, and detected using a horse anti-rabbit-Ig Cy5 antibody. Cells were examined under a Nikon A1 fluorescent microscope and micrographs taken using a 100x objective with a 2x digital zoom using the E2-C1 software package as described in the Materials and Methods.

Minimal inhibition of foreign DNA reporter gene expression by rhA3A and hA3A...
In a recent study, hA3A was shown to inhibit the expression of transfected plasmid DNA by deamination of the incoming plasmid DNA (Stenglein et al., 2010). Co-transfection of an eGFP reporter plasmid with hA3A in 293 cells resulted in approximately 20% reduction in the number of cells expressing eGFP signal 2 days. In these investigators’ model, the incoming DNA induces an interferon-induced up-regulation of hA3A expression, the cytidine residues are deaminated resulting in uracils, which are excised by UNG2. This will eventually result in cleavage and degradation of the DNA backbone (Stenglein et al., 2010). This phenomenon could significantly confound experiments that involve co-transfection of A3A with viral plasmid constructs.

Since our virus experiments involving rhA3A or hA3A lasted no longer than 48 h, we determined whether rhA3A, hA3A or rhA3G interfered with the expression of enhanced green fluorescent protein (eGFP) from a foreign plasmid. Plasmids expressing either rhA3A, hA3A or rhA3G were co-transfected into 293 cells 24 h prior to transfection with the plasmid expressing eGFP under the control of the CMV promoter. At 2 days post-transfection of eGFP, the cells were analyzed for fluorescence by flow cytometry. We detected 14.1, 14.5, 11.2, and 14.6% eGFP positive cells in the presence of empty vector, rhA3A, hA3A, or rhA3G, respectively (Figure 47A). These results indicate that hA3A had minimal impact (<20%) while rhA3A had no impact on reporter gene expression from foreign plasmids at 48 h.

**Human and rhesus macaque A3A does not inhibit SHIV or HIV-1 production in 293 cells**
Figure 47. Minimal restriction of foreign DNA by rhA3A and hA3A. Panel A. Expression constructs (3 µg) of HA-tagged rhA3A, hA3A, and rhA3G were first transfected into 293 cells for 24 h. At 24 h post-transfection, the 293 cells were re-transfected with a plasmid encoding a CMV-driven eGFP reporter and assayed for reporter expression. The cells were removed from the plate and fixed. All samples were analyzed by a flow cytometer and the mean fluorescence intensity (MFI) for eGFP positive cells was calculated. The MFI ratio and percentage of eGFP cells were calculated for each sample and the experiment was run in triplicate. All groups were compared to the pcDNA3.1(+) plus EGFP control using a Student’s t-test with p<0.05 considered significant. Panel B. SHIVΔVif production in the presence of rhA3A, hA3A and rhA3G. Cells were first transfected with 0.5 µg of plasmids expressing HA-tagged rhA3A, hA3A, rhA3G or empty vector. At 24 h post-transfection, the 293 cells were re-transfected with a plasmid containing the SHIVΔvif. Virus supernatants were harvested at 2 days post-transfection with the viral genome. Mean Gag p27 levels as measured using a commercially available ELISA kit (Zeptometrix) are shown. In all panels, error bars correspond to standard deviations from triplicate determinations. Panel C. HIV-1Δvif production in the presence of hA3A and rhA3A. Cells were first transfected with 0.5 µg of plasmids expressing HA-tagged rhA3A, hA3A, rhA3G, or empty vector. At 24 h post-transfection, the 293 cells were re-transfected with a 1 µg plasmid containing the HIV-1Δvif. Virus supernatants were harvested at 2 days post-transfection with the viral genome. Mean Gag p24 levels as measured
using a commercially available ELISA kit (Zeptometrix) are shown. In all panels, error bars correspond to standard deviations from triplicate determinations.
It can be argued that the slight inhibition of plasmid reporter gene expression by hA3A may be amplified with typically longer plasmids containing a full-length viral genome. Similarly, the absence of an effect by rhA3A in plasmid reporter expression may be enhanced in the context of a full-length viral molecular clone. We conjectured that if A3A can degrade viral infectious molecular clones, then virus production would be significantly inhibited. We therefore investigated whether virus production was inhibited in cells transfected with rhA3A or hA3A. 293 cells were transfected with the plasmids expressing hA3A, rhA3A, rhA3G or the empty plasmid 24 hours prior to transfection with plasmids with the SHIVΔvif or HIV-1Δvif constructs. At 2 days post-transfection with the viral genomes, the culture medium was assayed for either p27 (SHIV) or p24 (HIV-1). We did not observe a significant decrease in the level of SHIVΔvif p27 released into the culture medium in cells co-transfected with empty vector, rhA3A, hA3A, or rhA3G at 5 days post-transfection (Figure 47B). Similar findings were observed for HIV-1Δvif (Figure 47C). These data suggests that rhA3A and hA3A likely have no post-entry restriction activity against SHIV or HIV-1 and reinforce the conclusion that rhA3A and hA3A have minimal impact on plasmid DNA stability at 2 days post-transfection. Thus, our subsequent experiments, which involve analyzing cells within 48 h, should not be influenced by potential plasmid-specific restriction.

**Rhesus macaque A3A inhibits the infectivity of SHIVΔvif and HIV-1Δvif**

We next determined if rhA3A could inhibit the infectivity of SHIV and SHIVΔvif. For comparison, we also co-transfected these viral constructs with plasmids expressing
rhA3G, rhA3C, or hA3A. Virion infectivity was measured by taking the ratio of infectious titer in TZM-bl cells and comparing to the wild-type SHIV construct co-transfected with empty vector control. Our results indicate that rhA3A and rhA3G reduced the level of infectious SHIVΔvif nearly 20-fold compared to the parental SHIV, while hA3A reduced the infectivity of SHIVΔvif approximately 10-fold (Figure 48A). In contrast, rhA3C did not inhibit the level of infectious virus released from cells. SHIV virions produced in the presence of rhA3A, hA3A or rhA3G were purified and concentrated by ultracentrifugation and analyzed by Western blot. Consistent with the virion infectivity data, rhA3A was incorporated into SHIVΔvif and to a lesser extent in SHIV (Figure 48B). However, we were unable to detect hA3A in either SHIV or SHIVΔvif virions.

We also tested whether rhA3A had activity against HIV-1. HIV-1(NL4-3) and HIV-1Δvif constructs were co-transfected with plasmids expressing rhA3A, hA3A, or hA3G. Consistent with previous reports, hA3A did not inhibit infectivity of HIV-1Δvif virions, while hA3G reduced HIV-1Δvif virion infectivity by 10-fold. Surprisingly, rhA3A reduced HIV-1Δvif virion infectivity approximately 3-fold (Figure 48C). HIV-1 virions released from 293 cells expressing either hA3G, hA3A or rhA3A were partially purified by ultracentrifugation. The resulting virions were analyzed by Western blot using an anti-HA antibody. Consistent with previously published data, hA3A was detected in both HIV-1 and HIVΔvif virions. However, rhA3A was not detected in HIV-1 and HIVΔvif virions (Figure 48D). Together, these findings indicate that unlike hA3A, rhA3A reduces the infectivity of both SHIVΔvif and HIV-1Δvif.
Figure 48. Inhibition of SHIVΔVif and HIV-1ΔVif virion infectivity by rhA3A. Panel A. RhA3A restricts SHIVΔvif. 293 cells were co-transfected with 1 µg wild-type SHIV or SHIVΔvif molecular clones and 0.5 µg plasmids expressing either empty vector (pcDNA3.1), rhA3A, hA3A, or rhA3G. At 48 h, the culture medium was harvested and virion infectivity measured by taking the ratio of infectious viral titers on TZM-bl cells. Shown is the percentage virion infectivity with the empty plasmid control normalized to 100%. Panel B. 293 cells were transfected with wild-type SHIV or SHIVΔvif molecular clones and HA-tagged rhA3A, hA3A, or rhA3G. At 48 h, the culture medium was collected, clarified by low speed centrifugation, and concentrated by ultracentrifugation through a 20%/60% (w/v) step-gradient. Each sample was resuspended in 2x sample reducing buffer, boiled, and the proteins were separated by SDS-PAGE. The presence of rhA3G, hA3A, or rhA3A was detected by Western blot using an antibody directed against the HA-tag (Upper panel) and the presence of p27 using an antibody directed against p27 (Middle panel). The lower panel are 293 cells transfected with vectors expressing rhA3G, hA3A, rhA3A, or pcDNA3.1(+) vector in the absence of the viral genome. Panel C. RhA3A restricts HIV-1Δvif genomes were co-transfected with HA-tagged hA3A, hA3G, or rhA3A as described for the SHIV experiments in Panel A. Shown is the percentage virion infectivity with the empty plasmid control normalized to 100%. Panel D. Undetectable rhA3A in HIV-1Δvif virions. Virions were purified as described for the SHIV experiments in Panel B, and hA3G, hA3A, and rhA3A incorporation were detected by Western blot using an anti-HA
antibody, normalizing for Gag p24 levels. The lower panel are 293 cells transfected with vectors expressing rhA3G, hA3A, rhA3A, or pcDNA3.1(+) vector in the absence of the viral genome. Mean values in triplicate experiments are shown, and the statistical differences with the wild-type control were evaluated using a two-tailed Student’s t-test, with p<0.05, (▲) considered significant.
Deletion of amino acids 27-29 of rhA3A abolishes its anti-viral activity

Hominid and monkey A3As can be distinguished by a 3 amino-acid indel between residues 27 to 29. We hypothesize that this region may contribute to the differential activity of rhA3A and hA3A against SHIVΔvif and HIV-1Δvif. To test this hypothesis, we constructed a rhA3A mutant in which amino acids 27-29 (SVR) were deleted. The resulting mutant, rhA3AΔSVR, was found to be stable in cells (data not shown). Interestingly, rhA3AΔSVR was inactive against SHIVΔvif (Figure 49A) and HIV-1Δvif (Figure 49B). However, when the amino acids 27-29 (SVR) were introduced into hA3A, we restored the ability of hA3A to partially inhibit HIV-1Δvif (Figure 49B). These findings suggested that these monkey-specific A3A residues were critical for lentivirus inhibition, but was functionally lost in the hominid lineage.

RhA3A, but not hA3A, is sensitive to SIV and HIV-1 Vif-mediated degradation

We next determined the stability of rhA3A and hA3A in the presence of the SHIV, SHIVΔvif, HIV-1, or HIVΔvif genomes. Each full-length molecular clone was co-transfected into 293 cells along with plasmids expressing HA-tagged rhA3A, rhA3G, hA3A, or rhA3AΔSVR. At 24 h, the cells were lysed, proteins precipitated, and analyzed by Western blot using an antibody specific to the HA-tag. Our results show that rhA3G was degraded in the presence of the parental SHIV but not in cells co-transfected with the SHIVΔvif genome. In contrast, hA3A was stable in the presence of parental SHIV or SHIVΔvif (the apparent difference of the hA3A in lanes 4 and 5 is due to less protein
Figure 49. Deletion of three amino acids from the N-terminus region of rhA3A abolishes restriction. Panel A. RhA3AΔSVR is inactive against SHIVΔvif. 293 cells were co-transfected with (1.0 µg) SHIV and SHIVΔvif molecular clones with vectors expressing (0.5 µg) rhA3A, rhA3AΔSVR, rhA3G, or empty vector (pcDNA3.1) in a 12-well plate. At 48 h, the culture medium was harvested, assessed for p27 and the amount of infectious virus titrated onto TZM-bl cells. Panel B. RhA3AΔSVR is inactive against HIV-1Δvif. 293 cells were co-transfected with (1.0 µg) HIV-1 and HIV-1Δvif molecular clones with vectors expressing (0.5 µg) rhA3A, rhA3AΔSVR, hA3AΔSVR, hA3A, or empty vector (pcDNA3.1) in a 12-well plate. Virion infectivity was calculated normalizing to the wild-type HIV-1 empty vector control. In both panels, A and B, the assays were run at least in triplicate. Significance in the restriction was determined with respect to the wild-type empty vector control using a two-tailed Student’s t-test (p<0.05; ▲)
Figure 50. Sensitivity of rhA3A to lentivirus Vif-mediated degradation. Panel A. SHIV Vif mediates rhA3A degradation. SHIV and SHIVΔvif molecular clones were co-transfected with rhA3G, hA3A, rhA3A, or rhA3AΔSVR. At 24 h post-transfection, cells were lysed and proteins were precipitated using methanol. Proteins were separated on a 12% SDS-PAGE gel and probed using an antibody directed against the HA-tag. All samples were normalized to the amount of β-actin. Panel B. HIV-1 Vif mediates rhA3A degradation. HIV-1 and HIV-1Δvif molecular clones were co-transfected with rhA3G, hA3A, rhA3A, and rhA3AΔSVR. In both panels, blots were stripped and re-probed with an antibody against β-actin. All assays were performed twice.
loaded in lane 4). Unlike hA3A, rhA3A was degraded in the presence of the SHIV genome but stable in the presence of the SHIVΔvif genome (Figure 50A). From these results we conclude that rhA3A interacts with the SIV Vif protein. Finally, rhA3A appeared to be less stable in the presence of the HIV-1 genome and more stable in the presence of NL4-3Δvif (Figure 50B). Similar results were observed at 48 h post-transfection (data not shown). From these results we conclude that the SIV and HIV-1 Vif proteins may interact with rhA3A promote its rapid turnover.

**RhA3A causes low levels of G-to-A mutations in nascent SHIVΔVif reverse transcripts**

RhA3A incorporation into SHIVΔvif virions led to a significant decrease in infectivity (Figures 48B). Our previous study revealed that rhA3G inhibition of SHIVΔvif virions in target cells was accompanied by substantial G-to-A mutations (Schmitt et al., 2010). To determine whether rhA3A has a similar effect, we amplified a 300 bp region in *nef* from target TZM-bl cells and sequenced 15 independent clones.

The results indicate that while rhA3G caused 33 G-to-A mutations in 15 clones (300 bases sequenced in each clone or total of 4,500 bp) in the *nef* gene of SHIVΔvif, most of which (29) were in the context of 5’-TCT-3’ (Schmitt et al., 2010), while only 7 such mutations were detected with rhA3A (Figures 51B). As expected, the corresponding wild-type SHIV had significantly lower G-to-A mutation frequencies.
Figure 51. RhA3A induces low frequencies of G-to-A mutations in SHIVΔvif nascent reverse transcripts. TZM-bl cells were infected with DNase-I treated SHIV or SHIVΔvif virions and DNA was extracted after 24 h. A 300-bp segment of nef was amplified, cloned, and 15 independent clones were sequenced. Panel A. Moderate G-to-A mutations induced by rhA3A on SHIVΔvif. Panel B. Graph depicting the cumulative number of mutations from the 15 clones. RhA3A induced 5-fold lower G-to-A mutations than rhA3G against SHIVΔVif. Panel C. Lack of G-to-A mutations in wild-type SHIV in the presence of either rhA3A or rhA3G. Panel D. Graph of cumulative substitutions in wild type SHIV. In panels A and C, each mutation is denoted by a vertical line that is color coded with respect to the dinucleotide context: GA (red), GG (yellow), or non-G-to-A (black). In panels B and D, each bar is shaded according to the proportion of G-to-A substitutions that occurred in the context of GA (red), GG (yellow), or non-G-to-A (black).
Figure 52. RhA3A does not inhibit AAV-2 replication. Panel A. Low molecular weight DNA extracted from 293 cells transfected with recombinant AAV-2 plasmids and either empty vector (pcDNA3.1), hA3A, rhA3A, or rhA3G. A 1% Southern blot was run and hybridized with a $^{32}$P-labeled probe from SSV9 as previously described (Qiu and Pintel, 2002). The following AAV replication products are shown as RF (double recombinant form), RF (recombinant form), and ssDNA. These experiments were performed two times with identical results. Panel B. A Western blot was performed on an aliquot of the lysate to show that each A3 protein was expressed.
(Figures 53D). These results suggest that while rhA3A and rhA3G appear to have similar antiviral potency against SHIVΔvif (Figure 51D), rhA3A produced an approximately 5-fold decrease in the number of G-to-A mutations in SHIVΔvif compared to rhA3G.

**Rhesus A3A does not inhibit adeno-associated virus 2 (AAV-2)**

Previous studies have shown that hA3A can inhibit the replication of AAV-2 and autonomously replicating parvoviruses (Chen et al., 2006). We determined if rhA3A could inhibit the replication of AAV-2. 293 cells were transfected with vectors expressing hA3A, rhA3G, rhA3A or the empty vector and plasmids expressing AAV-2 and a helper plasmid. At 48 h, the cells were harvested, the extrachromosomal DNA isolated by Hirt extraction and analyzed using Southern blots for the presence of replicating AAV-2 DNA. The results, shown in Figure 52A, indicate that hA3A, but not rhA3G, inhibited AAV-2 replication, as previously reported (Narvaiza et al., 2009; Chen et al., 2006). In contrast, rhA3A did not significantly inhibit AAV-2 replication. A Western blot from an aliquot of the 293 lysates showed that the A3 proteins were similarly expressed (Figure 52B).

**RhA3A does not inhibit LINE-1 retrotransposition**
L1 elements are autonomous non-LTR retrotransposons that constitute about 17% of the human genome (Babushok and Kazazian, 2007). L1 elements, through the ORF2 gene product, also facilitate the retrotransposition of *Alu* elements, which are present in an additional 11% of the genome. These *in vitro* retrotransposition assays require the co-transfection of hA3A with L1 plasmids (Rangwala and Kazazian, 2009) and assaying after 3-4 days, a timepoint where significant hA3A inhibition of plasmid DNA may be observed (Stenglein et al., 2010).

To determine whether rhA3A could inhibit L1 retrotransposition, we used an L1-eGFP plasmid construct that contains a self-splicing intron in the opposite orientation within eGFP (pLRE3-EF1-mEGFPI) (Wissing et al., 2007). As a positive control, an intronless isogenic construct that encoded an intact eGFP (pLRE3-EF1-mEGFP) was used. L1 (1 µg) and A3A (200 ng) plasmids were co-transfected into 293T cells and the cells were harvested at 2 and 4 days. Similar to our earlier results (Figure 53B), at 2 days post-transfection, we observed minimal (22%) inhibition of the eGFP control plasmid, while rhA3A had no effect (data not shown). However, the L1 signals we obtained at 2 days were still very low, which is line with previous reported data (Kroutter et al., 2009). By 4 days post-transfection, we observed detectable L1 retrotransposition (Figure 53A). As expected, hA3A potently inhibited L1 retrotransposition. In contrast, rhA3A, as well as the rhA3A_SVR mutant, had no effect (Figure 53A). The effect of hA3A on L1 retrotransposition could be partially attributed to foreign DNA restriction, since at 4 days post-transfection, we observed a 2-fold reduction in eGFP signal from
the control plasmid (Figure 53A). However, it should be noted that substantial eGFP signals were still detected, in contrast to near-complete inhibition of L1 by hA3A (Figure 53A). Importantly, under conditions where we detected significant foreign DNA restriction by hA3A, no such activity was observed for rhA3A and rhA3A_SVR. Thus, we conclude that unlike hA3A, rhA3A does not restrict L1 retrotransposons and/or foreign DNA.

Discussion

SIV and SHIV infections of rhesus macaques have been used extensively as models for studying HIV pathogenesis and in vaccine development. The fact that the dose, timing and route of infection can be controlled also makes it ideal for investigating the earliest events following pathogenic lentivirus infection, including innate immune responses. Recent discoveries have shown that specific lentivirus genes could antagonize the effects of host innate restriction factors. Thus, the macaque/SIV or SHIV models provides a compelling system to explore how these restriction factors are regulated \textit{in vivo} following lentivirus infection.
Figure 53. RhA3A does not inhibit retrotransposition. One microgram of L1 constructs containing eGFP with or without the γ-globin intron in the opposite orientation and 200 ng of empty vector, hA3A, rhA3A, or rhA3AΔSVR plasmid were co-transfected into 293 T cells using Fugene 6 (Roche). At 4 days post-transfection, cells were harvested and analyzed for eGFP expression using a FACSCalibur machine (BD Biosciences) with 150,000 events. Panel A. Results of co-transfection of LINE 1 plasmid (with intron) with either the empty vector, hA3A, or rhA3A plasmids. To minimize autofluorescence background, the percentage of eGFP+ cells from the untransfected control was gated against an empty fluorescence channel (FL3). Panel B. Results of co-transfection of eGFP control plasmid (no introns) with either the empty vector, hA3A, or rhA3A plasmids.
Identification of rhA3A as a novel lentiviral restriction factor

The APOBEC3 family of deoxycytidine deaminases function as potent antiretroviral factors. Incorporation of select A3 proteins can inhibit reverse transcription and/or induce G-to-A hypermutation in nascent reverse transcripts in the next target cell (Bishop et al., 2004; Celico et al., 2006; Russell et al., 2009; Sova et al., 1993; von Schwedler et al., 1993; Zheng et al., 2005). Seven A3 genes are encoded in the human genome, but only a subset, hA3G, and to a lesser extent, hA3F and hA3C, are expressed in CD4+ T cells, inhibit HIV-1 and are counteracted by the HIV-1 encoded Vif protein (Koning et al., 2009; Refsland et al., 2010). There is uncertainty whether a similar complement of A3 genes are present in rhesus macaques, particularly since no rhA3A gene has yet to be identified. In this study, we cloned the rhA3A gene. This result is consistent with a recent study that revealed evolutionary conservation of A3A among hominids, Old World and New World monkeys (Bulliard et al., 2011). In addition, analysis of the rhesus macaque genome provided evidence of seven encoded A3 members that based on phylogenetic analyses are orthologous to the corresponding hA3 genes. These findings suggest that the seven rhA3 genes have been maintained after the evolutionary split between monkeys and humans. Thus, the vif genes of various SIV strains naturally infecting Old World monkey species have likely been co-evolving with the same set of A3 genes for millions of years.

Rhesus macaque A3G is a potent A3 deaminase against SIV, SHIV and HIV-1 (Schmitt et al., 2010; Virgen et al., 2007; Zennou et al., 2006). We therefore determined
the expression of the other rhesus macaque A3 genes in primary, activated CD4+ T cells, the main cellular targets of SIV and SHIV in vivo, in relation to rhA3G. Our results indicate that rhA3B mRNA is not expressed and that rhA3D mRNA is expressed at levels that are likely not biologically relevant in vivo. RhA3F is expressed at 2-fold lower levels than rhA3G, but rhA3F appears to be incorporated promiscuously into virions and is not as functional as rhA3G in side-by-side co-transfection studies (Schmitt et al., 2010). Only rhA3A, rhA3C and rhA3H are expressed at similar or higher levels than rhA3G. RhA3H has been previously shown to have potent antiretroviral activity against HIV-1 (OhAinle et al., 2008) and we found significant activity against SHIV (unpublished data). We also found no significant antiviral activity of rhA3C against SHIV. Surprisingly, rhA3A exhibited potent antiretroviral activity. Based on this data, we conclude that rhA3A, rhA3G, and rhA3H most likely are the dominant rhesus macaque APOBEC3 deaminases that could restrict SHIV or SIV in vivo.

**Insights on rhA3A restriction of lentiviruses**

Rhesus macaque A3G inhibits SHIVΔvif and HIV-1Δvif virion infectivity through a mechanism that involves virion incorporation. Virion incorporated rhA3G blocks reverse transcription and/or induces G-to-A hypermutation in the next target cell. Consistent with this mechanism, we readily detected rhA3G in SHIVΔvif and HIV-1Δvif virions, and observed substantial G-to-A mutations in reverse transcripts newly formed in target cells. Given these consistent phenotypes, we used rhA3G as a positive control for evaluating the antiretroviral activity of rhA3A. In this study, we found that rhA3A is as
potent as rhA3G in restricting virus infectivity. Both proteins inhibited SHIVΔvif infectivity by approximately 20-fold. In addition, rhA3A modestly inhibited HIVΔvif infectivity by about 3-fold. We found that in contrast to rhA3G, rhA3A was incorporated at low levels into SHIVΔvif virions and was undetectable in HIVΔvif virions (Figure 48). The results may be due to the sensitivity of the Western blots and that rhA3A is incorporated into virions at very low levels compared to rhA3G. This would also suggest that rhA3A enzymatic activity may be higher than rhA3G on a per-molar basis.

As a first step in gaining mechanistic insights on rhA3A-mediated restriction of SHIV and HIV-1, we evaluated whether two rhA3A domains are critical for restriction. Comparison of the amino acid sequences of human and rhesus macaque A3A reveals a monkey-specific insertion at position 27-29 (SVR) that is critical for SHIVΔvif and HIV-1ΔVif inhibition by rhA3A that was lost in the hominid lineage (Figure 41). In a recent modeling study of hA3A, it was shown that amino acids 29 and 30 likely form part of a polynucleotide accommodating groove near the active-site pocket (Bulliard et al., 2011). We observed that the rhA3A ΔSVR mutant is totally inactive, while remaining sensitive to SHIV and HIV-1 Vif-dependent degradation. In addition, we made just the opposite mutation, the addition of the SVR motif to hA3A. We found that this mutant gained partial activity against HIV-1Δvif but was not as active as rhA3A, suggesting that other amino acid substitutions are likely involved in this function. We also determined if the canonical cytidine deaminase Zn\(^{+2}\) binding domain (C/H-X-E-X\(_{23-28}\)-P-C-X\(_{2-4}\)-C) was critical for rhA3A anti-lentiviral activity. Mutation of the histidine at position 70 to glutamic acid resulted in no restriction of SHIVΔvif (unpublished data). Thus, the
monkey-specific residues SVR (27-29) and the canonical Zn$^{+2}$ binding domain are critical for rhA3A-mediated Δvif lentivirus restriction.

Human A3A is largely inactive against HIV-1ΔVif, but when targeted to the viral nucleocapsid, it has potent activity against HIV-1ΔVif (Goila-Gaur et al., 2007). Since hA3A and the C-terminal region of hA3G are derived from a common ancestor, the N-terminal domain of hA3G was fused to hA3A (hA3G-A3A). This hA3G-A3A fusion protein localized to the cytoplasm, incorporated into the viral nucleoprotein complexes, and inhibited HIV-1 replication (Goila-Gaur et al., 2007). Similarly, fusion of hA3A to the Vpr protein resulted in its incorporation into viral nucleoprotein complexes and inhibition of HIV-1 and SIV replication (Aguiar et al., 2008). These results suggested that the reason that hA3A could not restrict lentiviruses is that it could not be targeted to the virion core. While hA3A had no substantial activity against HIV-1ΔVif, hA3A surprisingly decreased SHIVΔVif virion infectivity by approximately 10-fold. Although we did not detect hA3A in virions; it is possible that like rhA3A, very little is incorporated into nascent virions and is below the limits of detection by Western blot analysis.

**Evidence for recent evolutionary gain of function of DNA restriction in hominids**

Previous studies have shown that while hA3A is inactive against HIV-1, it can block endogenous retroviral elements such as LTR-retrotransposons or Alu elements (Chen et al., 2006; Bogerd et al., 2006b). Human A3A has antiviral activity against autonomous (minute virus of mice; MVM) and non-autonomous (AAV-2) replicating
paroviruses and has editing activity against the human papillomavirus genome (Chen et al., 2006; Narvaiza et al., 2009; Vartanian et al., 2008). Finally, hA3A has been shown to counteract foreign DNA (Stenglein et al., 2010) and promote DNA damage (Landry et al., 2011). These studies suggest that hA3A biology is intimately linked to the restriction of double-stranded DNA elements. Surprisingly, we found no evidence for rhA3A restricting AAV-2, L1 elements and foreign DNA. While it can be argued that rhA3A may have activity against rhesus macaque AAV, no infectious macaque AAVs have been isolated to date. Thus, the DNA restriction properties of A3A may have been a recent evolutionary development in the hominid lineage.

The mammalian A3 locus encodes genes that have among the highest detectable evolutionary pressures known (Sawyer et al., 2004). This is thought to reflect a long-standing host-pathogen genetic conflict that eventually helps dictate current trends in cross-species transmission and viral host range. Our findings now enter A3A into this select group of genes, given the divergent virus specificities of the human and monkey homologues. RhA3A inhibits the SHIVΔvif and HIV-1ΔVif, but not the DNA virus AAV-2 and L1 retroelements, while hA3A inhibits SHIVΔvif, AAV-2 and L1 elements but not HIV-1Δvif. Thus, it would appear that rhA3A could inhibit a broader range of lentiviruses at the cost of having no activity against DNA viruses. On the other hand, hA3A appears to have acquired activity against DNA viruses, at the expense of losing broad potency against lentiviruses. The rhA3A mutant with amino acids 27-29 deleted (rhA3AΔSVR) was incapable of restricting SHIVΔvif and HIV-1Δvif. Thus, the divergent
viral specificities of hA3A and rhA3A may be dictated, at least in part, by an evolutionary switch that involved alterations in the polynucleotide-binding groove of A3A. The 81% identity between hA3A and rhA3A at the amino acid level should enable structure-based domain-swap experiments to determine which residues account for differential virus restriction. These types of studies should provide a biologically relevant system to interrogate which changes in the molecular determinants of nucleic acid specificity of the A3A proteins dictates the antiviral activity of A3A against lentiviruses, parvoviruses and L1 retroelements.

**Experimental Methods**

**Cells, viruses, and plasmids**

HeLa and 293 cells were used for transfections of vectors expressing various APOBEC3 proteins, full-length SHIV or HIV-1 (NL4-3). The TZM-bl cell line was used as an indicator cell line to measure the infectivity of viruses (Derdeyn et al., 2000; Platt et al., 2009; Takeuchi et al., 2008; Wei et al., 2002). Both cell lines was maintained in Dulbecco’s minimal essential medium (DMEM) with 10% fetal bovine serum (R10FBS), 10 mM Hepes buffer, pH 7.3, and 100 U/µg penicillin-streptomycin and 5 µg gentamicin. Rhesus macaque PBMCs were obtained from uninfected animals and isolated on Ficoll/Hypaque gradients. The derivation of SHIV$_{KU-2MC4}$ (a pathogenic molecular clone; referred to in the text as SHIV) and SHIV$_{VifSTOP}$ (referred to in the text as SHIV$\Delta$vif) has
been previously described (Liu et al., 1999; Schmitt et al., 2010). A plasmid with the genome of HIV-1 strain NL4-3 (referred to in the text as HIV-1) was used to construct a Δvif version (referred to in text as HIV-1Δvif) (pNL4-3; NIH AIDS Research and Reference Reagent Program). This plasmid was digested with PflM1, phenol: chloroform extracted, and ethanol precipitated. The 5'-protruding ends were filled using DNA Polymerase I Large Fragment (Klenow; Promega). The resulting Klenow fragment reaction produced blunt ended fragments that were purified and re-ligated. The resulting plasmid was sequenced and found to contain a deletion in pNL4-3 vif between base pairs 5301 to 5308. Therefore, only the first 79 amino acids of pNL4-3 Vif could be expressed. The plasmids pcDNA3.1(+)HA-rhA3G and pcDNA3.1(+)HA-hA3A were kindly provided by Nathaniel Landau (New York University School of Medicine). A plasmid expressing HA-rhA3C was constructed in this laboratory and cloned into the pCruz-HA vector (Santa Cruz Biotechnology) using restriction sites Sca I and Kpn I. A plasmid expressing pcDNA3.1(+)HA-hA3G used in these studies was provided by the NIH AIDS Reference and Reagent Program.

**Molecular cloning of rhA3A**

Total RNA from mitogen-activated rhesus macaque PBMCs was isolated using the RNeasy Kit (QIAGEN; Valencia, CA). Three hundred ng of sample was subjected to random hexamer priming using the RT\(^2\) EZ first strand kit (SA Biosciences), and 4 µl of this cDNA was subjected to a 40 µl PCR reaction consisting of 1x Sweet PCR Mix (SA Biosciences), 12.5 pmol of rhA3A.F (5'-GACAAGCACATGGACGGCAG) and rhA3A.R
(5'-CATCCTTCAGTTTCCCTGAT) primers, designed based on the first and last exons of human A3A (hA3A). Amplification conditions included 95°C for 15 min hot-start, followed by 40 cycles at 94°C 30 s, 50°C 30 s and 72°C 3 min in a PE 9700 machine (Perkin Elmer). A dominant PCR amplicon of 600 bp was subcloned into the TOPO-TA vector (Invitrogen) and sequenced. The sequence for rhesus macaque A3A (rhA3A) was submitted to GenBank with Accession number JF831054. An expression construct containing rhA3A linked to an N-terminal HA-tag was synthesized (GenScript). The HA-rhA3A gene was subcloned into pcDNA3.1(+) using restriction sites Kpn I and Eco RV sites of pcDNA3.1(+) and sequence-confirmed.

Positioning of rhA3 genes in the rhesus macaque genome

Analysis of the Apobec3 loci was performed using NCBI Build 1.2 version of the macaque chromosome 10 sequence. RhA3C (Accession # EU381233), rhA3G (XM_001094452), rhA3F (DQ514917) and rhA3H (DQ507277) cDNA sequences were subjected to BLAST analysis against the macaque genome, and the corresponding coordinates were obtained. To characterize the intronic sequence flanked by exons 1 and 2 of rhA3A, exon 1 (5'-AGAAGAGACAAGCACATGGAC) and exon 2 (5'-GCTCCACCTCGTAGCACAA) specific primers were used to directly amplify rhesus macaque genomic DNA. The resulting PCR amplicon was directly sequenced (Figure 42; see text for details). RhA3B and rhA3DE sequences were predicted based on the genome sequences upstream of rhA3C and between rhA3C and rhA3F, respectively, using the geneid program (http://genome.crg.es/geneid.html) (Figures 43 and 44).
**Phylogenetic analysis of rhApobec3A**

The A3A amino acid sequences from various primate species were obtained from (Bulliard et al. 2011). Sequences were aligned using ClustalX using the Gonnet series, and neighbor-joining trees were constructed with 1000 subreplicates, correcting for multiple substitutions and excluding gaps. The C-terminal half of rhA3G was used as outgroup. A3 amino acid sequences from human, mouse, rat and rhesus macaque were likewise aligned using Clustal X, after splitting the sequences of rodent A3 and primate A3B, A3DE, A3F and A3G into N-terminal and C-terminal halves. Neighbor-joining trees were performed unrooted, with 1000 subreplicates and excluding gaps.

**Expression of rhA3A in rhesus macaque tissues**

A rhesus macaque that euthanized and perfused with 2 liters of Ringer’s saline to remove contaminating blood was used to examine the expression of rhA3A mRNA in different tissues (visceral organs and brain). RhA3A was amplified from 30 µg of each visceral and CNS tissue using the RNAEasy kit (Qiagen; Valencia, CA). RNA was digested with 1 U/µl DNase-I (Fermentas) for 30 min. RT-PCR was performed using the Titan One RT-PCR kit (Roche). Each reaction used 100 ng of total RNA and was amplified using oligonucleotide primers specific for rhA3A. The oligonucleotides used for the reverse transcriptase reaction and first round PCR were 5’-GGACGGCAGCCAGCATCCAGGCCCAG-3’ (sense) and 5’-CGTAGGTCATGATGGAGACTTGGGC -3’ (antisense), which are complementary to
bases 3-29 and 451-475, respectively. The nested oligonucleotides used were 5’-CCAGGCCAGACACTTGATGGATCC-3’ (sense) and 5’-CCGCAGCGTTCGCAGTGCCTCCTGATACAGG-3’ (antisense), which are complementary to bases 20-44 and 411-441, respectively. As a control, GAPDH was amplified to verify the integrity of the RNA using oligonucleotides 5’-GCCATCACTGCCACCCAG-3’ (sense) and 5’-GCCACATACCAGAAATGAGC-3’ (antisense). The nested oligonucleotides used were 5’-CCTCCGGGAAACTGTGGC-3’ (sense) and 5’-CGTTGAGGGCAATGCCAG-3’ (antisense). The reactions were performed in an ABI 2720 Thermal Cycler using the following thermal profile: 42°C 30 min for 1 cycle; 94°C 2 min for 1 cycle, 94°C 30 s, 55°C 30 s, and 68°C 45 s for 10 cycles; 94°C for 30 s, 55°C 30 s, and 68°C 2 min for 25 cycles. One µl of the initial reaction mixture used for nested PCR using rTaq (Takara) and performed with the following thermal cycle profile: 95°C for 1 min, 48°C 2 min, and 72°C 3 min for 35 cycles. The 450 bp amplicon was visualized in a 1.5% agarose gel, excised and gel purified (Qiagen; Valencia, CA) and directly sequenced.

**Expression of rhA3 genes in activated CD4^+ T cells**

Rhesus PBMCs were obtained from three donors by Ficoll separation. 5’ 10⁶ cells were seeded into each well of a 6-well plate containing 3 ml RPMI medium (Mediatech) with 10% Fetal Bovine Serum (Gemini Biosciences). Half of the cultured PBMCs were activated with 2.5 µg/ml Staphylococcal Enterotoxin B (SEB; Toxin Technology) overnight. The following day, the PBMCs were stained with CD3-V450
(clone SP34-2), CD4-APC-H7 (L200) and CD25-PerCP-Cy5.5 (M-A251) (BD Biosciences) and analyzed in an LSR-II flow cytometer (BD Biosciences). Histogram plots to highlight CD25 expression in CD4+ T cells were constructed using the FlowJo software (Treestar). CD4+ T cells were purified from activated PBMCs by negative selection with magnetic beads (Miltenyi Biotec). Total RNA was extracted from CD4+ T cells using the RNAEasy kit, and cDNA was synthesized by random hexamer priming using the RT² EZ first strand kit (SA Biosciences). The cDNA was used for triplicate real-time PCR evaluations (25 µl) with optimized A3-specific primers and a housekeeping gene, GAPDH, in a BioRad real-time PCR machine. Briefly, 5 µl of cDNA was mixed with 12.5 µl of 1´ Universal Master Mix (Applied Biosystems), 10 pmol of primers and 5 pmol of probe and subjected to 40 cycles of denaturation and annealing/elongation. The A3-specific primers, probes and cycling conditions are listed in Table VIII. Cloned cDNA sequences were used as standards and for evaluating cross-reactivity (Table VIII). The A3-specific assays had a limit of detection of 5 copies per reaction, and standard curves had r² values >98%.

**Subcellular localization of rhA3A**

For intracellular localization of rhA3A, 293 cells were plated onto 13 mm cover slips in a 6 well plate and transiently co-transfected with vectors expressing HA-tagged hA3A, rhA3G or rhA3A and one expressing a eGFP tagged nuclear marker using a polyethylenimine (PEI) transfection reagent (ExGen™ 500, MBI Fermentas) according to the manufacturer’s instructions. Cultures were maintained for 36-48 h before being
were fixed in 2% paraformaldehyde (4°C) for 10 min and washed twice for 5 min in 1x PBS (pH7.4). Cover slips were incubated with a rabbit polyclonal anti-HA antibody (HA-probe, Santa Cruz) for 1 h at ambient temperature in 1x PBS plus 1% BSA. Cover slips were washed twice for 5 min in PBS and reacted with a Cy-5 conjugated secondary antibody (Abcam) 1x PBS plus 1% BSA for 30 min. The cover slips were washed twice for 5 min in 1x PBS and mounted in glycerol containing mounting media (Slowfade antifade solution A, Invitrogen). A Nikon A1 confocal microscope was used to collect 100x images with a 2x digital zoom, using EZ-C1 software. The pinhole was set to large for all wavelengths used with Cy5 and eGFP excited using a 638 nm and 488 nm diode lasers, respectively. Images were collected using a 670 nm filter for Cy5 and 525/25 nm filter for eGFP.

**Foreign DNA restriction by rhA3A**

293 cells were seeded into 6-well tissue culture plates 24 h prior to transfection. Cells were first transfected with 3 mg of either HA-rhA3A, HA-hA3A, HA-rhA3G HA-rhA3ADSVR or pcDNA3.1(+) vector using a polyethylenimine transfection reagent (ExGen™ 500, MBI Fermentas) according to the manufacturer’s instructions. After 24 h, the cells were re-transfected with 3 mg of a vector expressing eGFP. Cells were incubated at 37°C in 5% CO₂ atmosphere for either 2 or 5 days. The cells were removed from the plate using Ca²⁺/Mg²⁺-free PBS containing 10 mM EDTA. Cells were then fixed in 2% paraformaldehyde, for 5 min. The cells were washed twice in 1x PBS plus and analyzed using an LSRII flow cytometer. The mean fluorescence intensity
(MFI) for eGFP positive cells was calculated. The MFI ratio and percentage of eGFP positive cells was calculated for each sample. Normalized ratios from three separate experiments were averaged and the standard error calculated. All groups were compared to pcDNA3.1(+) plus eGFP control using a Student’s t-test with $p<0.05$ considered significant.

**Production of HIV-1ΔVif and SHIVΔVif virions in the presence of A3A**

293 cells were seeded into a 12-well tissue culture plate 24 h prior to transfection. Cells were first transfected with 1 mg of HA-tagged rhA3A, hA3A, rhA3G or pcDNA3.1(+) vector using PEI (Fermentas). After 24 h, cells were transfected with 1 mg of SHIVΔvif or HIV-1Δvif plasmid. Cells were incubated at 37°C in 5% CO$_2$ atmosphere for 48 h. Supernatants were collected and cellular debris removed by low speed centrifugation. The cells were lysed in 500 ml of RIPA (50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 0.5% deoxycholate, 0.2% SDS, 10 mM EDTA) and the nuclei were removed through high-speed centrifugation. The amount of Gag p27 or p24 present in the viral supernatants was measured using a commercially available p27 and p24 ELISA kits (Zeptometrix). The experiment was run at least three separate times, normalized to the empty vector control samples. Differences between mean percentages were calculated using a two-tailed Student’s t-test with $p<0.05$ considered significant.

**Inhibition of SHIV and HIV-1 virion infectivity by A3A**
SHIV, SHIVΔVif, HIV-1 or HIVΔVif infectious molecular clones (3 µg) were co-transfected (1.5 µg) with plasmids expressing rhA3A, rhA3AΔSVR, rhA3G, rhA3C, hA3A, hA3A+SVR, or hA3G using PEI (Fermentas) in a 6-well plate. At 48 h post-transfection, the culture medium was harvested and clarified by low speed centrifugation. Equivalent amounts of p27 were serially diluted using 10-fold dilutions from 10^1 to 10^6 and used to inoculate TZM-bl cells. At 48 h post-inoculation, the media was removed, cells washed with PBS and the monolayer fixed using 1% formaldehyde-0.2% glutaraldehyde in PBS. The cells were washed and incubated in a solution for 2 h at 37°C containing 4 mM potassium ferrocyanide, 4 mM potassium ferricyanide, 4 mM magnesium chloride, and 0.4 mg X-gal per ml. The reaction was stopped and the infectious units (IU) per ml were calculated (Derdeyn et al., 2000; Wei et al., 2002).

**Site-directed mutagenesis**

Mutations introduced into all plasmids were accomplished using a QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer’s protocol. All plasmid inserts were sequenced to ensure the validity of the mutations and that no other mutations were introduced during the cloning process.

**Virion incorporation assays**

SHIV, SHIVΔVif, HIV-1 or HIVΔVif infectious molecular clones (3 µg) were co-transfected with 1.5 µg of plasmids expressing either rhA3A, rhA3G, hA3G or hA3A proteins into 293 cells using PEI reagent (Ex-Gen™ 500, Fermentas) in each well of a
6-well plate. At 48 h, virus supernatants were harvested and clarified by low speed centrifugation. The clarified supernatant was ultracentrifuged to pellet virions (SW41 rotor, 247,000xg, 1 h, 4°C). The pellet was resuspended in PBS and layered on a 20%/60% sucrose step gradient and again subjected to ultracentrifugation (SW55Ti, 247,000´g, 1 h) at 4°C. The interface (containing virions) was harvested, pelleted again by ultracentrifugation as described above, and resuspended in 150 µl of PBS. An aliquot was saved to determine the p27 or p24 content by antigen capture assay (Zeptometrix). The remaining sample was boiled in 2´ sample reducing buffer. Equivalent amounts of p27 or p24 were loaded on a 12% SDS-PAGE gel. A3 proteins were detected by Western blotting using an anti-HA antibody (HA-probe; Santa Cruz). Blots were placed in stripping buffer (25 mM glycine, pH 2.0 and 1% SDS) and reprobed using either rabbit polyclonal antibody specific for p27 or a mouse monoclonal antibody to p24 (NIH AIDS Research and Reference Reagent Program).

**Stability of rhA3A in the presence of SHIV and HIV-1 genomes**

Infectious molecular clones of HIV-1, HIV-1Δvif or SHIV and SHIVΔvif were co-transfected in a 2:1 ratio with rhA3A, rhA3AΔSVR, rhA3G, hA3G, or hA3A using PEI transfection reagent (Ex-Gen500; Fermentas) into a 12-well plate. At 24 h post-transfection, the supernatant was removed, the cells were harvested and lysed using RIPA buffer. Following lysis, the nuclei were removed by centrifugation at 14,000 rpm for 15 min in a microfuge at 4°C. The protein was precipitated using methanol, resuspended in 2x sample reducing buffer, and boiled for 5 min. Proteins were
separated on a 12% SDS-PAGE gel and probed using commercially available rabbit polyclonal HA antibody (HA-probe, Santa Cruz). All samples were normalized to the same amount of β-actin protein using a mouse monoclonal antibody specific for β-actin (AC15; Novous Biologicals).

**Hypermutation assays in the presence of rhA3A**

To determine whether virion-incorporated rhA3A can induce G-to-A mutations in nascent reverse transcripts, 293 cells were co-transfected with vectors containing the genomes SHIV or SHIVΔVif and vectors expressing rhA3G (positive control), rhA3C (negative control) or rhA3A using PEI (Fermentas). At 24 h post-transfection, cells were washed and fresh DMEM was added. At 48 h, the supernatant containing virus was subjected to ultracentrifugation through a 20% sucrose cushion. The resulting supernatant was DNase-I-treated (Fermentas) at 37°C for 30 min to minimize plasmid carry-over. Infectious titers in the supernatants were measured on TZM-bl cells. In addition, at 24 h post-infection of TZM-bl cells, total cellular DNA was harvested and extracted using the DNeasy kit (Qiagen). The DNA was used in a nested DNA PCR to amplify a 300 base pair fragment of nef. The PCR reaction was carried out using rTaq and the manufacturer’s instructions (Takara). The oligonucleotides employed during in the first round PCR to amplify SIV nef were 5’-GGTGGAGCTATTCCCAGGAGG-3’ (sense) and 5’-GTCTTTGGACTGTAATTAAATCCC-3’ (anti-sense). One µl of the first PCR product was added to a nested reaction. The oligonucleotides used during the nested PCR reaction were 5’-CCATGAGCGGTCCAGGCACTGAG-3’ (sense)
and 5'-CCTCCCAGTCCCCCCT TTTC-3' (anti-sense). The PCR reactions were performed using an ABI 2720 Thermal Cycler with the following thermal profile: 95°C for 2 min, 1 cycle; 95°C for 30 s, 55°C for 30 s, 65°C for 2 min, for 35 cycles; 65°C for 7 min. The PCR products were separated by electrophoresis, isolated, purified, sequenced and sub-cloned into pGEM-TEasy (Promega). Fifteen independent clones were sequenced and assessed for each mutant SHIV as described above for each condition.

**Adeno-associated virus-2 (AAV-2) replication assay**

The ability of rhA3A to inhibit the replication of AAV2 was assessed. 293 cells were transfected using LipoD 293 transfection reagent (SignaGen Labs; Rockville, MD) into 60 mm dishes in duplicate with 1 mg SSV9 (an AAV2 infectious clone, psub201) (Qiu et al., 2002), 2mg pHelper (Stratagene, La Jolla, CA) (Xiao et al., 1998), and 1 mg of an A3 expression vector, rhA3G, hA3A, rhA3A, or as a negative control pBluescript SK (+) (Stratagene, La Jolla, CA). Forty-eight h post-transfection, the cells were harvested and resuspended in PBS. The supernatant was discarded and Hirt DNA was extracted using Hirt solution (10 mM Tris, 10 mM EDTA, pH 7.5, 0.6% SDS) at room temperature for 10 min followed by 5 M NaCl at 4°C overnight. The mixture was centrifuged at 14,000 rpm at 4°C for 20 min. Following centrifugation, the pellet was discarded and the supernatant was treated with Proteinase K (50 mg/ml) at 37°C for 1 h. The low molecular weight (LMW) DNA was extracted twice with saturated phenol and once with chloroform/isoamyl alcohol. The resulting DNA was ethanol precipitated and column purified using the Qiagen gel extraction kit (Qiagen). The LMW DNA was run on
a 1% agarose gel, transferred to nitrocellulose and analyzed by Southern hybridization using a $^{32}$P-labeled probe consisting of a fragment representing nucleotides 184-4490 ($Xba$ I digestion) from SSV9 as previously described (Qiu et al., 2002). To ensure that all samples were expressing the desired A3 proteins, aliquots of harvested cells utilized in the Southern blot were lysed in RIPA buffer. The nuclei were removed and the protein was precipitated using methanol, resuspended in 2X sample reducing buffer, and boiled for 5 min. Proteins were separated on a 12% SDS-PAGE gel and probed using a rabbit polyclonal anti-HA antibody (HA-probe, Santa Cruz).

L1 retrotransposition assay

The engineered L1 enhanced green fluorescent protein (EGFP) reporter (Ostertag et al., 2000) was cloned into the $pLRE3-EF1-mEGFPI$, which contains a retrotransposition competent L1 element (LRE3) (Brouha et al., 2003) under the control of an EF-1α promoter in addition to the internal 5' UTR promoter. The enhanced GFP (eGFP) retrotransposition indicator cassette is under the control of an ubiquitin promoter (UBC), and the SV40 late polyadenylation signal. The construct was cloned into $pBSKS-II+$ (Stratagene). The positive control $pLRE3-EF1-mEGFP(\Delta\text{intron})$ is similar to the $pLRE3-EF1-mEGFPI$ but lacks the intron in the $mEGFPI$ indicator cassette and therefore serves as a positive control for transfection efficiency. One µg of L1 construct and 200 ng of empty vector, hA3A, rhA3A or rhA3AΔSVR plasmid was co-transfected into 293T cells using Fugene 6 (Roche). All transfections were performed in triplicate.
Cells were harvested at 2 and 4 days and analyzed for eGFP expression using a FACSCalibur machine (BD Biosciences), collecting 150,000 events.

Previous research has shown that human APOBEC3A (hA3A) lacks antiviral activity against HIV-1, but inhibits the replication of adeno-associated virus 2 (AAV-2), intracisternal A particles (IAP), and long interspersed nuclear element 1 (L1) (Bogerd et al., 2006a,b; Chen et al., 2006; Muckenfuss et al., 2006). Although hA3A contains one CDA, the deaminase activity of hA3A is not required for its inhibitory role against AAV or retrotransposons, suggesting that a deaminase-independent mechanism of inhibition is occurring (Narvaiza et al., 2009). Human A3A also contains significant homology to the N-terminal half of hA3G, does not associate with the viral nucleocapsid complex (NPC) of HIV-1 but is readily incorporated into virions (Goila-Gaur et al., 2007).

SIV and SHIV have been invaluable models for studying various aspects of HIV-1 pathogenesis. Sequencing of the rhesus macaque genome has led to the identification of rhesus homologues of A3B, A3C, A3D, A3F, A3G and A3H. Here we describe the identification of an A3A homologue in rhesus macaques (rhA3A) and present evidence that the human and macaque Apobec3 genes are orthologous. We show that the rhA3A protein is expressed well in activated macaque CD4+ T cells and other rhesus macaque tissues, is incorporated into virions, is degraded by both HIV-1 and SHIV in a Vif dependent manner, and restricts the replication of SHIV (expressing the SIVmac239 Vif) and HIV-1.
IX. Conclusion

In the last thirty years, HIV-1 has become one of the most devastating infectious diseases in humans. Developing countries display the greatest HIV-1-related morbidity and mortality, which is primarily due to the fact that only 10% of infected patients receive HAART therapy. HAART therapy prolongs the lifespan of HIV-1 infected patients by decreasing CD4+ T cell loss and progression to AIDS. Presently, HAART therapy does not cure the patient, and there is an increased frequency of multidrug resistant HIV-1 strains. Therefore, new small molecule inhibitors or a vaccine to prevent infection or reduce transmission of HIV-1 is necessary to minimize the expansion of this epidemic.

In addition to humans, HIV-1 only infects chimpanzees and a practical alternative to studying the role of HIV-1 accessory genes in viral pathogenesis is to use macaques infected with non-human primate lentiviruses (Heeney et al., 2006). Amino acid substitutions in these proteins can be studied using Asian (rhesus or pig-tailed) macaques infected with either SIV or SHIV to study disease progression, test antiretroviral therapy, and possible vaccine strategies (Ambrose et al., 2007). While these models have provided us with insight into the role of these genes in pathogenesis, they also have short comings such as: 1) the diversity of viral protein sequences between different primate lentiviral lineages; 2) the evaluation of viral proteins in replication or as antiviral targets in vivo; 3) the testing of vaccine epitopes only identified in HIV-1; and 4) anti-retroviral drug efficacy and efficiency (Thippeshappa et al., 2011). Our laboratory has studied the role of HIV-1 Vpu and its various domains in CD4+ T cell loss, virus release, and pathogenesis in macaques using the chimeric SHIVs: SHIV\textsubscript{KU-1bMC33} and
SHIV_{KU-2MC4} (Stephens et al., 2002; Singh et al., 2003; Hout et al., 2005; 2006; Hill et al., 2008).

**Highly Conserved Domains in the C-Terminus of Primate Lentivirus Vif Proteins**

Cellular proteins termed restriction factors have been identified and found to function in preventing the spread of retroviruses and endogenous mobile genetic elements (Goff et al., 2004; Malim et al., 2008; Albin et al., 2010). As a consequence, many retroviruses have evolved mechanisms to counter this restriction. The HIV-1 accessory protein Vif is required for lentivirus infectivity of specific cell types (except EIAV); therefore, HIV-1 virions deficient in the Vif protein cannot sustain a productive infection in either primary CD_{4}^{+} T cells, macrophages, and many established CD_{4}^{+} T cell lines (H9, C8166, CEMx174) (Gabuzda et al., 1992; Madani et al., 1998; Simon et al., 1998; Sova et al., 1993). These nonpermissive cells express an antiviral restriction factor known as APOBEC3G (A3G) (Sheehy et al., 2002). Vif binds APOBEC3 through non-linear binding sites in the N-terminus. Following the Vif-APOBEC3G interaction, Vif recruits the E3 ubiquitin ligase complex, which consists of elongin B (EloB), elongin C (EloC), Cullin-5 (Cul5), and Rbx-1 (Kao et al., 2003; Mehle et al., 2004a; 2006; Yu et al., 2003). This complex induces polyubiquitination of APOBEC3G and subsequent degradation by the 26S proteosome (Conticello et al., 2009; Marin et al., 2003; Mehle et al., 2004; Sheehy et al., 2003; Stopak et al., 2003; Yu et al., 2004). There are two highly conserved domains near the carboxy-terminus of Vif that have been implicated in the recruitment of the Cul5-E3 ubiquitin ligase machinery. One of the domains identified is
the SLQ(Y/F)LA motif, which was shown to mediate the binding of Vif to EloC (Mehle et al., 2004b; Yu et al., 2003; Yu et al., 2004). Therefore, one of the goals of this dissertation was to study highly conserved domains of primate lentivirus Vif in vivo, and determine their roles in pathogenicity.

Previous in vitro research has shown that mutation of the serine, leucine, and glutamine prevent the association of Elongin BC complex subsequently disruption of the proteasomal targeting and degradation of the A3 proteins resulting in a non-productive infection (Fang et al., 2007; Mehle et al., 2004; Yu et al., 2003; Yu et al., 2004). Mutations in this motif have also been shown to increase A3 incorporation into virions resulting G-to-A hypermutation (Kobayashi et al., 2005). The importance of the SLQ(Y/F)LA domain of Vif in its interaction with the Elongin BC complex in viral pathogenesis is essential to understanding the specific motifs required for the degradation of A3 proteins. This data can potentially assist in the development of small molecule inhibitors or attenuated vif vaccines targeting both conserved and functional domains of Vif that limit viral replication and increase the host immune response. Therefore, the first study focused on determining if specific mutations in the SLQYLA motif of SIV\textsubscript{mac}239 Vif were critical for viral pathogenesis in rhesus macaques.

We mutated two residues of the highly conserved SLQYLA motif in SHIV, which expresses a SIV\textsubscript{mac}239 Vif. The serine at amino acid position 147 and the leucine at amino acid position 148 were mutated to alanine residues resulting in a virus called SHIV\textsubscript{VifAAQYLA}, which was used to inoculate three rhesus macaques. These macaques developed no significant loss in circulating CD4+ T cells, displayed 100-fold lower viral
loads, no histological lesions, and fewer visceral tissues positive for viral gag DNA or RNA in two of the three macaques, when compared to macaques inoculated with SHIV\textsubscript{KU-1bMC33}. In two of the three macaques (RAK10 and RCS10), specific tissues such as the mesenteric, axillary, and inguinal lymph nodes, tonsil, and thymus, were positive for both viral RNA and DNA. The presence of viral RNA suggests that these macaques have active viral replication in the primary and secondary lymphoid organs and that the virus may have persisted in these tissue populations. Recently, Refsland and colleagues published studies regarding the profiling of the A3 mRNA repertoire in human lymphocytes and tissues using quantitative PCR methods (Refsland et al., 2010). From their studies they profiled the expression of A3 mRNA in twenty tissues such as adipose, bladder, brain, cervix, colon, esophagus, heart, kidney, liver, lung, ovary, placenta, prostate, skeletal muscle, small intestine, spleen, testes, thymus, thyroid, and the trachea. Interestingly, the lung, adipose, spleen, bladder, thymus, heart, and cervix were among the tissues that showed a broad expression of A3 mRNA (Refsland et al., 2010). Previously, our laboratory used immunohistochemical analysis to determined whether A3G was expressed in the brain and kidney. It was found that A3G expression in brain was limited to the pyramidal neurons in the gray matter of the cerebral and cerebellar cortices and restricted to the proximal convoluted tubules of kidney in the pigtailed macaque (Hill et al., 2006; 2007). However, since we are using the rhesus macaque as a model of SHIV infection, both the visceral tissues and specific regions of the brain need to be identified for the expression of rhesus A3A/B/C/D/F/G/H mRNA. No studies to date have examined this, and experimentally it would allow laboratories using
the rhesus macaque as a nonhuman primate to study the restricted replication patterns or identify reservoirs of \textit{vif} attenuated SHIV replication. The replication of virus in specific tissues may be explained by the expression patterns of A3 mRNA in those tissues.

In HIV-1 Vif deficient viruses, A3 proteins are incorporated into the virion, which are effectively introduced into the target cell upon the next round of infection (Soros et al., 2007). The A3 proteins can then mediate extensive dC to dU mutations of the minus-strand of viral DNA during reverse transcription (Suspene et al., 2004; Yu et al., 2004). These mutated reverse transcripts can then be destroyed either by the DNA base repair enzymes such as uracil DNA glycosylase or apurinic-apyrimidinic endonuclease or survive to serve as a template for plus-strand synthesis resulting in the accumulation of G-to-A substitutions (Harris et al., 2003; Mangeat et al., 2003). If these G-to-A substitutions (plus strand) exceed 10% of all of the viral dG residues, this is referred to a “hypermutation” of the viral genome (Harris et al., 2003; Lecossier et al., 2003; Mangeat et al., 2003; Zhang et al., 2003). These mutations can result in alteration of viral open reading frames, decreased tRNALys3 primer annealing to the primer binding site on the genomic RNA, and mutation of essential viral genes required for a productive HIV-1 infection (Harris et al., 2003; Lecossier et al., 2003; Mangeat et al., 2003; Suspene et al., 2004; Yu et al., 2004; Zhang et al., 2003). Often the G-to-A mutations (plus strand) occur within highly polarized mutational gradients in the viral genome from 5’ to 3’ due to the time that these regions of the minus-strand of DNA remain single-stranded (Suspene et al., 2006; Yu et al., 2004). Our sequencing results of the thymus, mesenteric, axillary,
and inguinal lymph nodes of macaques inoculated with SHIV\textsubscript{VifAAQYLA} indicated that a definite 5’ to 3’ (minus strand) gradient of G-to-A substitutions occurred in the viral genome. The \textit{vif}, \textit{vpu}, \textit{nef}, and \textit{env} genes were sequenced and assessed in these tissues resulting in an overall gradient in two of the three macaques analyzed, \textit{nef}>\textit{env}>\textit{vpu}>\textit{vif}. The exception to this gradient was macaque RPL10 who displayed an overall high number of G-to-A (plus strand) transitions in all four genes. Over 85% of these mutations were in the dinucleotide sequence context of 5’-TC, we concluded that an A3 protein besides A3G was involved in the cytidine deamination mutational patterns observed. It is important to note that not all lymphoid tissues showed G-to-A substitutions above background levels. Since A3F, A3D, A3B, A3C, and A3H can display this mutational pattern, it is hard to decipher if one or many of these A3 proteins were involved in the cytidine deamination. Therefore, tissue specific expression patterns of various A3 mRNA would help us further determine which of those four tissues we analyzed expressed higher levels of one A3 protein over another. Recently, we discovered the A3A protein in the rhesus macaque and looked at an expression panel of \textit{rhA3} mRNA in Staphylococcal Enterotoxin B (SEB) activated CD4\textsuperscript{+} T cells. Using quantitative real-time RT-PCR, the expression level of A3A, A3C, and A3H relative to A3G were 1.0-1.5-fold similar in magnitude, while A3F and A3D were expressed on average 2.4- and 13.6-fold lower in magnitude than A3G, respectively. Therefore, one of these three A3 proteins (A3A, A3C, A3G, or A3H) is most likely causing cytidine deamination in the viral DNA with a dinucleotide context of 5’-TC (minus strand).
Interestingly, when assessing sequences from SHIV_{VifAAQYLA}^-inoculated macaque PBMC DNA at weeks 1, 3, 4, 8, 12, and 28 weeks post-inoculation, two of the three macaques (RPL10 and RAK10) displayed a G-to-A mutation at four weeks. This mutation resulted in the mutation of the site-directed alanine at amino acid position 147 to a threonine in the 147AAQYLA 152 (to 147TAQYLA 152) motif of Vif. The mutation of amino acid residue 147 demonstrates that not all G-to-A substitutions are detrimental to the viral genome. While we do not know why RAK10 and RPL10 selected for this mutation or how pathogenicity of the virus is affected, we do know that it did not alter viral burdens or levels of circulating CD4+ T cells in these macaques beyond four weeks post-inoculation. One possible theory could be that the A147T mutation could have permitted a more productive SHIV infection in these macaques resulting in a better antigenic stimulus. Macaques RAK10 and RPL10 both had higher viral burdens and developed immunoprecipitating antibodies similar to the positive control; whereas, macaque RCS10 retained the A147A mutation, had lower viral burdens, and significantly less immunoprecipitating antibodies. More research is needed to assess the replication fitness, pathogenicity, and structural/mechanistic implications of the A147T mutation selected from these rhesus macaques.

To extend our in vivo studies and further cripple the SHIV_{VifAAQYLA} mutant, we constructed a SHIV mutant that expresses a Vif protein with the first five amino acids of the SLQYLA substituted for alanine residues (SHIV_{Vif5A}). In vitro SHIV_{Vif5A} in the presence of rhesus A3G was readily incorporated into virions, displayed a 1,000-fold decrease in viral infectivity, revealed a high frequency of G-to-A substitutions in the
dinucleotide context of 5’-TC (minus strand) in the nef gene. Macaque inoculated with SHIV\textsubscript{Vif5A} displayed a further reduction in plasma viral loads compared to the SHIV\textsubscript{VifAAQYLA} inoculated macaques. The initial burst of replication occurred at one-week post-inoculation and occasionally detected in the plasma throughout the duration of infection, but maintained levels of circulating CD4\textsuperscript{+} T cells at pre-inoculation levels. However, sequences were obtained from PBMC only during the first three weeks of infection. At necropsy, the viral RNA copy number in the visceral and central nervous system tissues could not be quantified due to the limit of detection for this assay being ~100 copies. Since we were unable to sequence the nef gene in the PBMC throughout the course of infection or at necropsy, we were unable to assess the number of G-to-A substitutions in the viral genome. Also, these macaques produced immunoprecipitating antibodies by 12 weeks post-inoculation that waned by necropsy. Since viral RNA was not detected in the visceral tissues of macaques inoculated with SHIV\textsubscript{Vif5A} at necropsy, the infectious virus could have been readily cleared by the macaque, and viral proteins may not have been produced to provide an antigenic stimulus for an increased immune response against the infection. We were unable to detect any viral loads in the visceral organs or the CNS tissue in macaques inoculated with SHIV\textsubscript{Vif5A}. These results could be due to SHIV\textsubscript{Vif5A} replicating initially in cell types that express a wide range of APOBEC3 proteins that were incorporated into the virion and exerted an antiviral affect on the next target cell. During the next round of replication, the viral genome could have been highly mutated with G-to-A substitutions, and the RNA polymerase II could not transcribe the viral RNA efficiently. Also, there may have been tissue or cellular reservoirs where
APOBEC3 proteins are not expressed and there is no restriction on the replication of SHIV<sub>Vif5A</sub>. As described above, further studies will be needed to determine if cell populations exist that do or do not express APOBEC3 proteins. These results show that primate lentiviruses expressing a Vif protein with targeted mutations in a highly conserved domain (SLQ(Y/F)LA) can be controlled by the host, are more detrimental to the virus <em>in vivo</em>, and that this domain is critical to Vif function <em>in vivo</em>.

In addition, another highly conserved motif is the zinc-binding motif H-X<sub>5</sub>-C-X<sub>17</sub>-C-X<sub>16</sub>-H (HCCH) motif, which has been shown to interact with Cul5 (Luo et al., 2005; Mehle et al., 2004; Yu et al., 2004). Previous tissue culture studies showed that mutation of the two cysteine residues or chelation of zinc inhibited proper folding of HIV-1 Vif and prevented the recruitment of Cul5 resulting in the inability of Vif to recruit the E3 ubiquitin ligase complex (Mehle et al., 2006; Paul et al., 2006; Xiao et al., 2006; 2007a; 2007b). For this study, we constructed a SHIV where four amino acids in the HCCH motif of the Vif protein were mutated to alanine residues, resulting in the virus SHIV<sub>VifHCCH(-)</sub>. <em>In vitro</em> SHIV<sub>VifHCCH(-)</sub> in the presence of rhesus A3G, readily incorporated rhA3G, showed a 100-fold decrease in viral titer, and displayed an increase in the number of G-to-A substitutions in the <em>nef</em> gene in the context of 5'-TC and not 5'-CC (minus strand). Macaques inoculated with SHIV<sub>VifHCCH(-)</sub> showed slightly higher viral loads than SHIV<sub>Vif5A</sub> at one week post-inoculation. However, by three weeks post-inoculation, the virus was effectively controlled by the macaque with bursts of replication at weeks 6, 20 and 24. These bursts of replication correlated with G-to-A substitutions in the dinucleotide context of 5'TC (minus strand) that we observed in the <em>vif</em> gene amplified
from PBMC DNA. Immunoprecipitating antibodies were detected at 12 weeks post-inoculation, and were non-existent by necropsy. We also assessed the distribution of viral loads in the visceral tissues of macaques inoculated with SHIV\textsubscript{VifHCCH(-)} relative to SHIV\textsubscript{KU-2MC4}. We found that there was a significant decrease in viral loads of macaques infected with SHIV\textsubscript{VifHCCH(-)}, when compared to macaques inoculated with SHIV\textsubscript{KU-2MC4}, suggesting that SHIV\textsubscript{VifHCCH(-)} infection is significantly less widespread. As described above, these results suggest that SHIV\textsubscript{VifHCCH(-)} may replicate in the macaque in a cell type-specific manner depending on APOBEC3 expression. Also, the viral RNA detected in visceral tissues of macaques inoculated with SHIV\textsubscript{VifHCCH(-)} at necropsy may not have been translated into viral proteins due to excessive G-to-A mutation; therefore, the infection may have been cleared by the macaque because no antigenic stimulation through viral proteins could occur. Also, due to hypermutation viral quasi-species may exist allowing for the co-existence of functional, deleted, or truncated vif genes which can determine the degree of viral attenuation depending on the site and extension of the deletion (Rangel et al., 2009).

Using lentiviruses as a live attenuated vaccine candidate has been involved in the removal of “non-essential” viral genes such as vpr, vpx, nef, or vpu, which results in low levels of viral replication, the accumulation of compensatory mutations in other genes, and/or recombination or reversion to a more pathogenic phenotype. These live attenuated virus vaccines are usually based on non-virulent cloned viral variants or clones engineered with deletions in accessory genes (Abel et al., 2003; Almond and Scott, 1999; Daniel et al., 1992; Lohman et al., 1994). Following vaccination, challenge
with pathogenic SIV or SHIV should protect the host by suppressing the challenge viral loads, increasing the survival of the host, and preservation of the CD4+ T cell population. The live attenuated vaccine approach is not always successful in protecting against SIV and the length of time between vaccination and challenge seems to be a variable that influences efficacy; however, many of the factors required to determine the protective efficacy have not yet been identified (Clements et al., 1995; Marthas et al., 1990; Wyand et al., 1996). While the immune correlates for HIV-1 infection are still unknown, a vaccine needs to ensure relative safety and induce a potent protective immune response in the host (Wyand et al., 1996). Vaccines consisting of the deletion or inactivation of one or more viral genes on HIV-1 or SIV seem to be the most effective (Koff et al., 2006).

Macaque studies examining vif-deleted lentiviruses are limited, but all of the studies thus far have shown that Vif is required for viral replication in vivo (Desrosiers et al., 1998; Gabuzda et al., 1994; Harmache et al., 1996; Inoshima et al., 1996; Kristbjornsdottir et al., 2004; Lockridge et al., 1999; Sparger et al., 2008). In a previous study the roles of various accessory proteins of SIV$_{mac239}$, including the $vif$ gene, were assessed. Rhesus macaques inoculated with SIV$_{mac239 \Delta vif}$ maintained low viral loads, did not show any viral DNA in PBMC at weeks 2 or 16, and did not produce antibodies to SIV viral proteins such as vpx or vpr (Desrosiers et al., 1998). In a second study, investigators showed that a $vif$-deleted proviral DNA vaccine inoculated into rhesus macaques was immunogenic displaying a virus-specific T cell proliferative response even though the virus was severely attenuated. Even though the virus was administered as a booster at 6 and 22 weeks post-inoculation, the vaccine was not protective against
challenge with SIVmac251 (Sparger et al., 2008). Our studies showed that both SHIVVifHCCH(-) and SHIVVif5A were effectively controlled by rhesus macaques during the primary phase of infection. This suggests that viruses containing mutations in the conserved domains of vif may be useful to create a vaccine candidate for HIV-1. Since these macaques were singly inoculated with each virus and developed an immunoprecipitating antibody response at 12 weeks post-inoculation, it would be of interest to determine if multiple inoculations would result in a stronger immune response that could protect against challenge with a pathogenic SHIV. Since the Vif protein is required for replication of HIV-1 (SHIV/SIV) in CD4+ T cells and macrophages, mutations targeted in conserved motifs of Vif could permit limited replication and effective immune responses (both humoral and cell mediated) thus representing a novel means of producing a live attenuated lentivirus vaccine candidate. Furthermore, more research is required to determine if mutations in other conserved motifs in Vif my permit more replication to improve immune responses prior to being controlled by the host, and whether they would protect against challenge with pathogenic virus. For instance, while the C-terminal region of Vif has been studied in this dissertation, the N-terminal region of Vif has not been assessed in vivo (Figure 13). The N-terminal region of HIV-1 consists of various non-linear binding sites for A3 proteins. Since there is only 50% amino acid homology between HIV-1 SIVmac239, it would be of interest to determine the amino acids in SIVmac239 that are responsible for binding A3 proteins. It is important to functionally understand and identify these domains to determine if they are highly conserved, like “SLIK” or “YxxL.” This would provide more conserved binding sites in Vif
that could be used as attenuated viruses to boost the host immune response and protect against challenge.

**Identification of Rhesus APOBEC3A**

Human APOBEC3 (hA3) genes are a large family of tissue specific cytidine deaminases that are tandemly arrayed on chromosome 22 (Jarmuz et al., 2002). Human A3B, A3C, A3D, A3G, and A3F have been shown to inhibit the replication of HIV-1Δvif and SIVmac239Δvif to various extents (Dang et al., 2006; 2008; Doehle et al., 2005; Mariani et al., 2003; Wiegand et al., 2004; Yang et al., 2007; Yu et al., 2004b; Zennou and Bieniasz, 2006; Zheng et al., 2004). While hA3A lacks antiviral activity against HIV-1 and HIV-1Δvif, it inhibits the replication of adeno-associated virus 2 (AAV2), intracisternal A particles (IAP), long interspersed nuclear element 1 (LINE-1, L1), and Alu, through deaminase-independent mechanisms (Bogerd et al., 2006a; 2006b; Chen et al., 2006; Goila-Gaur et al., 2007; Muckenfuss et al., 2006). Even though hA3A lacks antiviral activity, it is still readily incorporated into the virions in the absence of the viral nucleoprotein complex (NPC) in the presence of a functional Vif. Human A3A contains significant homology to the C-terminal region of hA3G. Therefore, when the N-terminal half of hA3G was fused to hA3A, the A3G-A3A chimera was effectively packaged into only HIV-1Δvif virions by associating with the NPC, showed cytoplasmic localization similar to hA3G, and exhibited a strong antiviral activity against HIV-1Δvif (Goila-Gaur et al., 2007).
While hA3A has been extensively studied, no rhesus A3A (rhA3A) counterpart had been cloned from the rhesus macaque, suggesting that rhA3A may have been deleted in the rhesus macaque genome (OhAinle et al., 2008; Virgen and Hatzioannou, 2007). However, a recent publication showed that human A3A contains orthologs in eleven different non-human primate species such as owl and African green monkeys (Bulliard et al., 2011). Therefore, we were interested in the isolating and characterizing the antiviral activities of rhesus A3A against HIV-1 and SIV_{mac}239 Vif. Based on the rhesus macaque genome sequence, we were able to clone rhesus A3A and determine that this gene is tandemly arrayed with the other six rhesus A3 genes on chromosome 10 in the rhesus macaque. We determined that the rhA3A cloned amplicon revealed 82% nucleotide sequence identity with hA3A, and that the rhesus sequence together with Asian macaques (“Old World” monkeys) contained a primate lineage-specific insertion at amino acid positions 27 to 30 in A3A genes. We further classified the rhA3A gene in a phylogenetic analysis determining that the rhesus macaque and human A3A genes are orthologous.

We found that rhesus A3A is widely expressed in both the visceral and central nervous system tissues of the rhesus macaque, is expressed at high levels in Staphylococcal Enterotoxin B (SEB) activated CD4^{+} T cells from three donor macaques, and degraded in the presence of both HIV-1 and SIV_{mac}239 genomes. While it is thought that localization can partially explain the differences in antiviral activity against retroviruses and retrotransposons, this was not the case (Bogerd et al., 2006a; 2006b; Chen et al., 2006; Goila-Gaur et al., 2007; Muckenfuss et al., 2006). In our study, rhA3A
displayed nucleocytoplasmic localization similar to hA3A. However, previous reports suggest that hA3A can inhibit AAV-2, IAP, L1, and Alu. In our studies rhA3A did not restrict the replication of AAV-2 or L1 retrotransposition (Bogerd et al., 2006a; 2006b; Chen et al., 2006; Goila-Gaur et al., 2007; Muckenfuss et al., 2006).

Recent research by Stenglein and colleagues suggests that hA3A in vitro can inhibit the expression of transfected plasmid DNA through deaminase-dependent mechanisms, while genomic DNA is not affected. These authors suggest that foreign DNA restriction is a conserved innate immune defense mechanism predominantly in monocytes, macrophages, and neutrophils (Stenglein et al., 2010; 25, 29). Therefore, genetic engineering, gene therapy, and DNA vaccines may be influenced by this foreign DNA defense mechanism effecting the efficiency and fidelity of these treatments (Stenglein et al., 2010). However, our studies using an eGFP expression plasmid or full-length viral molecular clones co-transfected with either hA3A or rhA3A did not show significant differences at 48 hours (the length of most of our tissue culture experiments). Human A3A had minimal impact (<20%), while rhA3A did not have an impact on either reporter gene expression or full-length viral DNA molecular clone. Also, when we examined the release of virus from these cells, rhA3A and hA3A had no post-entry restriction activity against SHIV or HIV-1. Therefore, our studies should have minimal impact on plasmid DNA stability throughout the duration of our experiments. Also, the DNA restriction properties of hA3A may be a recent evolutionary development in the hominid lineage.
Similar to rhA3G, rhA3A was a potent inhibitor of SHIVΔvif by reducing the levels of infectious virus by 20-fold. Interestingly, we also found that hA3A effectively inhibited SHIVΔvif to a lesser extent by reducing the levels of infectious virus by 10-fold. Unlike SHIVΔvif, rhA3A only reduced the infectivity of HIV-1Δvif by 3-fold. Therefore, for the first time we were able to show an evolutionary switch in the antiviral activities of primate A3A by showing that rhA3A can inhibit lentiviral replication. Previous research suggests that virion incorporation of A3 proteins is necessary for A3 to exhibit its antiviral deamination-dependent effects on HIV-1 in the next target cell (Gladdis et al., 2003; Sheehy et al., 2002; Suspene et al., 2004). However, we were only able to detect small amounts of rhA3A incorporation into SHIV virions and unable to detect its incorporation into HIV-1 virions. Also, even though hA3A inhibited SHIVΔvif, hA3A was not detected in SHIV virions. This may be due to the sensitivity our virion incorporation technique suggesting that small amounts of rhA3A or hA3A are incorporated into SHIV virions. Using in vitro hypermutation assays, we were able to observe G-to-A (plus strand) substitutions in nef of SHIVΔvif in the presence of rhA3A. However, the frequency of G-to-A substitutions observed with rhA3A was significantly less (a 5-fold decrease) than the G-to-A substitutions observed with rhA3G. Therefore, rhA3A may be potently restricting SHIVΔvif through both deaminase-dependent and deaminase-independent mechanisms.

Since the hominid and monkey A3A’s are distinguished by a three amino acid indel (insertion and/or deletion) between residues 27 to 29 (SVR), we sought to determine if this region is necessary for the differential antiviral activity of hA3A and
rhA3A. When the “SVR” amino acids were removed from rhA3A, SHIVΔvif and HIV-1Δvif were no longer restricted by rhA3AΔSVR but still remained sensitive to SHIV and HIV-1 Vif-mediated degradation. However, when the “SVR” amino acids were inserted into hA3A, hA3A+SVR gained partial activity against HIV-1Δvif. This demonstrates that these amino acid residues are critical for rhA3A lentiviral inhibition, but they may have been functionally lost during the evolutionary split between humans and monkeys. A recent publication by Bulliard and colleagues modeled the structure of hA3A based upon the known crystal structure of the hA3G C-terminus (Holden et al., 2008; Shandilya et al., 2010). This analysis found that the hA3A structure tightly correlates with loop 1 of the predicted DNA docking/editing groove (polynucleotide-accommodating groove) in the C-terminus of hA3G. Based upon their work, the “SVR” amino acids point towards the protein core providing stabilization to rhA3A and form a groove that accommodate a single-stranded DNA molecule for cytidine deamination (Bulliard et al., 2010).

Previous research has determined that hA3A inhibits the replication of AAV-2, autonomously replicating paroviruses, intracisternal A particles (IAP), long interspersed nuclear element 1 (LINE-1, L1), and Alu, through deaminase-independent mechanisms (Bogerd et al., 2006a; 2006b; Chen et al., 2006; Goila-Gaur et al., 2007; Muckenfuss et al., 2006). As previously reported, we found that hA3A can inhibit AAV-2, while rhA3G cannot (Chen et al., 2006; Narvaiza et al., 2009). However, we showed that rhA3A was unable to restrict the replication of AAV-2 and LINE-1 elements. This necessity may not be required as no infectious AAVs have been isolated from rhesus macaques.
Based on these results, further studies are needed to examine the restriction patterns of primate A3A genes from both a structural and evolutionary perspective. Therefore, experiments designed to determine why hA3A does not restrict HIV-1Δvif and why rhA3A potently restricts of SHIVΔvif and to a lesser extent HIV-1Δvif are necessary. Using the C-terminal crystal structure of hA3G and determining a molecular model obtained using the interactive SWISS-MODEL program of both hA3A and rhA3A, we will be able to identify residues near the polynucleotide binding groove or residues that are solvent exposed, suggesting that they may have the ability to interact with single-stranded DNA (Figure 53) (Bulliard et al., 2010; Holden et al., 2008; Shandilya et al., 2010). In order to study this we would need to construct chimeric domain swapping protein mutants, this will allow us to examine which changes in molecular determinants of nucleic acid specificity of A3A proteins dictates the antiviral activity of A3A against lentiviruses or the restriction of AAV-2 replication and LINE-1 retrotransposition. These domain-swapping proteins would be constructed as follows: AC-Loop1 (accommodating loop 1; D35G, L26I, S27, V28, R29, Q33K), AC-Loop 2, AC-Loop 3 (C59H, K61Q, V64L, P66I, D69F, C72R, V74A), AC-Loop 5 (R106S, R107W), and α-helix 4 (Q140K, R144Q, T145M) (Bulliard et al., 2010; Holden et al., 2008; Shandilya et al., 2010). The AC-Loop 1 is thought to have proximity to the active center and bind the target DNA for deamination, while α-helix 4 may be important for A3A specificity and polynucleotide binding. These studies will allow us to understand the structural basis involved in restriction of lentivirus and parvovirus replication as well as the inhibition of LINE-1 retrotransposition. Also, determining if the biological properties of the restriction by A3A...
proteins from Old World and New World monkeys would be interesting. This is because New World macaques have not been evolutionarily driven by SIV like Old World macaques. Looking at these A3A proteins would help us determine the potential driving forces for A3A evolution. For example, sequence analysis shows that A3A proteins from various Old World monkeys contain variable amino acids in the “SVR” motif. Therefore, it would be interesting to see if other mutations from various Old World monkeys would still inhibit lentiviral replication.

Viral protein X (Vpx) is uniquely encoded by HIV-2 and SIV (SIV_{mac}/SIV_{sm}) and has been found to bind hA3A to counteract its antiviral effects (Berger et al., 2010; Horton et al., 1994; Wolfrun et al., 2007). Vpx is required for the infection of human monocytes and monocyte-derived cells such as macrophages and dendritic cells in vitro but dispensable for the infection of primary lymphocytes and cell lines (Berger et al., 2010; Cheng et al., 2008; Fletcher et al., 1996; Sharova et al., 2008; Wolfrum et al., 2007). SIV_{mac} Vpx partially binds hA3A and decreases its steady-state levels within the cell through degradation during the early phase of myeloid cell infection allowing for the replication of SIV_{mac}. A single-point mutation in SIV_{mac} Vpx, H82A, prevents binding hA3A allowing hA3A to abrogate the infection of monocytes (Berger et al., 2010; Berger et al., 2011). Therefore, it would be interesting to design experiments to determine if SIV_{mac} Vpx binds rhesus A3A allowing rhA3A to abrogate the infection of monocytes. With the domain swap mutant proteins described above, we could determine which domains of human and rhesus A3A are required for this interaction. This could reveal
another novel role for rhA3A by inhibiting lentiviral replication and shed light on the importance of myeloid cells during primate lentiviral pathogenesis.
Figure 53: Structural differences between human and rhesus A3A. Panel A. The predicted human A3A structure based off of the x-ray crystal structure of the C-terminal half of human A3G (Bulliard et al., 2010; Holden et al., 2008). Panel B. The predicted structure of rhesus A3A based on the SWISS-MODEL program and the x-ray crystal structure of the C-terminal half of human A3G (Dr. Xiaojiang Chen, University of Southern California).
A.

B.
X. References


2 Abel, K. et al. Simian-human immunodeficiency virus SHIV89.6-induced protection against intravaginal challenge with pathogenic SIVmac239 is independent of the route of immunization and is associated with a combination of cytotoxic T-lymphocyte and alpha interferon responses. Journal of virology 77, 3099-3118 (2003).


21 Arens, M. et al. Alterations in spliced and unspliced HIV-1-specific RNA detection in


45 Bennasser, Y. & Bahraoui, E. HIV-1 Tat protein induces interleukin-10 in human


60 Blanc, D., Patience, C., Schulz, T. F., Weiss, R. & Spire, B. Transcomplementation of


67 Bonvin, M. *et al.* Interferon-inducible expression of APOBEC3 editing enzymes in


74 Breuer, S. *et al.* Biochemical indication for myristoylation-dependent conformational


89 Chang, H. C., Samaniego, F., Nair, B. C., Buonaguro, L. & Ensoli, B. HIV-1 Tat protein


145 Derdeyn, C. A. et al. Sensitivity of human immunodeficiency virus type 1 to the fusion inhibitor T-20 is modulated by coreceptor specificity defined by the V3 loop of gp120. *Journal of virology* 74, 8358-8367 (2000).


150 Doehle, B. P., Schafer, A. & Cullen, B. R. Human APOBEC3B is a potent inhibitor of HIV-1 infectivity and is resistant to HIV-1 Vif. *Virology* 339, 281-288,


Fang, Q. et al. [The apoptosis-inducing activity of human selenoprotein P shorter...


174 Fonseca, V. A. & Kahn, S. E. Commentary: Of HOPEs and DREAMs: the quest to


183 Fujita, M. *et al.* Amino acid residues 88 and 89 in the central hydrophilic region of


Gao, F. et al. Evidence of two distinct subsubtypes within the HIV-1 subtype A radiation. 


Ghys, P. D. et al. The associations between cervicovaginal HIV shedding, sexually transmitted diseases and immunosuppression in female sex workers in Abidjan, Cote


207 Gooch, B. D. & Cullen, B. R. Functional domain organization of human APOBEC3G.


Guo, F. *et al.* The interaction of APOBEC3G with human immunodeficiency virus type 1


230 Hahn, B. H., Shaw, G. M., De Cock, K. M. & Sharp, P. M. AIDS as a zoonosis: scientific


239 Harris, R. S. & Liddament, M. T. Retroviral restriction by APOBEC proteins. *Nature*

240 Harris, R. S., Petersen-Mahrt, S. K. & Neuberger, M. S. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Molecular cell* 10, 1247-1253 (2002).


254 Hill, M. S. *et al.* APOBEC3G expression is restricted to neurons in the brains of pigtailed macaques. *AIDS research and human retroviruses* **22**, 541-550,


275 Huthoff, H. & Malim, M. H. Identification of amino acid residues in APOBEC3G required
for regulation by human immunodeficiency virus type 1 Vif and Virion encapsidation.


283 Jabbar, M. A. & Nayak, D. P. Intracellular interaction of human immunodeficiency virus type 1 (ARV-2) envelope glycoprotein gp160 with CD4 blocks the movement and


287 Janini, M., Rogers, M., Birx, D. R. & McCutchan, F. E. Human immunodeficiency virus type 1 DNA sequences genetically damaged by hypermutation are often abundant in patient peripheral blood mononuclear cells and may be generated during near-simultaneous infection and activation of CD4(+) T cells. Journal of virology 75, 7973-7986 (2001).


Jordan, A., Defechereux, P. & Verdin, E. The site of HIV-1 integration in the human
genome determines basal transcriptional activity and response to Tat transactivation.


306 Kahn, S. M. Making technology familiar: orthodox Jews and infertility support, advice,


---


328 Kijak, G. H. *et al.* Variable contexts and levels of hypermutation in HIV-1 proviral


Kozak, S. L., Marin, M., Rose, K. M., Bystrom, C. \& Kabat, D. The anti-HIV-1 editing


Langlois, M. A., Beale, R. C., Conticello, S. G. & Neuberger, M. S. Mutational comparison of the single-domained APOBEC3C and double-domained APOBEC3F/G


Liu, B., Yu, X., Luo, K., Yu, Y. & Yu, X. F. Influence of primate lentiviral Vif and


379 Lundquist, C. A., Tobiume, M., Zhou, J., Unutmaz, D. & Aiken, C. Nef-mediated downregulation of CD4 enhances human immunodeficiency virus type 1 replication in


388


Malim, M. H. & Cullen, B. R. HIV-1 structural gene expression requires the binding of


401 Margottin, F. *et al.* Interaction between the cytoplasmic domains of HIV-1 Vpu and CD4: role of Vpu residues involved in CD4 interaction and in vitro CD4 degradation. *Virology*


Masuyama, N. et al. HM1.24 is internalized from lipid rafts by clathrin-mediated endocytosis through interaction with alpha-adaptin. *The Journal of biological chemistry*


McCormick-Davis, C., Dalton, S. B., Singh, D. K. & Stephens, E. B. Comparison of Vpu sequences from diverse geographical isolates of HIV type 1 identifies the presence of


423 Melo, F. & Feytmans, E. Assessing protein structures with a non-local atomic interaction


430 Miller, M. A. *et al.* Alterations in cell membrane permeability by the lentivirus lytic peptide (LLP-1) of HIV-1 transmembrane protein. *Virology* **196**, 89-100,


Miyagi, E., Andrew, A. J., Kao, S. & Strebel, K. Vpu enhances HIV-1 virus release in the


Ogert, R. A. *et al.* N-linked glycosylation sites adjacent to and within the V1/V2 and the V3 loops of dualtropic human immunodeficiency virus type 1 isolate DH12 gp120 affect coreceptor usage and cellular tropism. *Journal of virology* **75**, 5998-6006,


Pagans, S. *et al.* SIRT1 regulates HIV transcription via Tat deacetylation. *PLoS Biology*


Perron, M. J. *et al.* The human TRIM5alpha restriction factor mediates accelerated


hypermutation gradient: purifying selection at multiple steps during HIV-1 replication results in levels of G-to-A mutations that are high in DNA, intermediate in cellular viral RNA, and low in virion RNA. *Retrovirology* 6, 16, doi:10.1186/1742-4690-6-16 (2009).


Schrofelbauer, B., Chen, D. & Landau, N. R. A single amino acid of APOBEC3G controls its species-specific interaction with virion infectivity factor (Vif). *Proceedings of


568 Schwartz, O., Marechal, V., Le Gall, S., Lemonnier, F. & Heard, J. M. Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein.


Shindo, K. *et al.* The enzymatic activity of CEM15/Apobec-3G is essential for the


419


620 Stremlau, M. *et al.* Specific recognition and accelerated uncoating of retroviral capsids


626 Svarovskaia, E. S. *et al.* Human apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) is incorporated into HIV-1 virions through interactions


of Sciences of the United States of America 86, 5163-5167 (1989).


Van Heuverswyn, F. *et al.* Human immunodeficiency viruses: SIV infection in wild


657 Venkatesh, L. K. & Chinnadurai, G. Mutants in a conserved region near the carboxy-


Wei, B. L. et al. Activation of p21-activated kinase 2 by human immunodeficiency virus type 1 Nef induces merlin phosphorylation. *Journal of virology* **79**, 14976-14980,


Williams, S. A. *et al.* NF-kappaB p50 promotes HIV latency through HDAC recruitment


693 Wray, V. *et al.* Solution structure of the hydrophilic region of HIV-1 encoded virus protein U (Vpu) by CD and 1H NMR spectroscopy. *International journal of peptide and protein


Xiao, Z. et al. Assembly of HIV-1 Vif-Cul5 E3 ubiquitin ligase through a novel zinc-


Yu, Q. *et al.* APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *The Journal of biological chemistry* **279**, 53379-


Zauli, G. et al. HIV-1 Tat protein down-regulates CREB transcription factor expression in PC12 neuronal cells through a phosphatidylinositol 3-kinase/AKT/cyclic nucleoside phosphodiesterase pathway. The FASEB journal : official publication of the Federation


728 Zheng, L., Yang, Y. D., Lu, G. C. & Salvato, M. S. Extracellular HIV Tat and Tat cysteine rich peptide increase CCR5 expression in monocytes. Journal of Zhejiang University.


XI. Appendix


