# STUDIES ON THE ORIGINS OF HIV-1 MUTATION AND GENETIC DIVERSITY

A DISSERTATION SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF THE UNIVERSITY OF MINNESOTA

Colleen M. Holtz

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Advisor: Louis M. Mansky

December 2013

**ACKNOWLEDGEMENTS**

I would like to thank:

My husband,

My family,

My advisor,

My committee,

My fellow lab members,

The DDS/Ph.D program,

MinnCResT

**Abstract**

A fundamental biological property of retroviruses and RNA viruses is their ability to rapidly mutate and evolve. The ability of these viruses to generate high levels of genetic diversity during replication has clearly had a profound impact on their ability to maintain their niche in nature, and to rapidly adapt to changing environmental conditions or opportunities to expand their host range. Previous reports with HIV-1 have indicated that the cell type in which HIV-1 replicates does not have a profound impact on HIV-1 diversity. However, due to differences in dNTP pool levels and expression levels of HIV-1 DNA editing enzymes, the hypothesis that cell type does influence the diversity of HIV-1 populations was formulated. To test this, a panel of relevant cell types (i.e., CEM-GFP, U373-MAGI, 293T, and SupT1) was analyzed for their ability to influence HIV-1 mutant rate and spectra. No differences were observed in overall mutation rate, but intriguingly, cell type differences impacted HIV-1 mutation spectra. These observations represent the first description of significant differences in HIV-1 mutation spectra observed in different cell types in the absence of changes in the viral mutation rate and, imply that such differences could have a profound impact on HIV-1 pathogenesis, immune evasion, and drug resistance.

The most common mutation type that arises during HIV-1 replication is transition mutations, particularly G-to-A mutations. Apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like 3 (APOBEC3) proteins create G-to-A mutations at specific cytosine dinucleotides. In order to better define the locations of APOBEC3G (A3G)-mediated G-to-A mutations, we tested the hypothesis that sequence context and DNA secondary structure influence the creation of A3G-mediated G-to-A mutations. Single-

stranded DNA secondary structure as well as the bases directly 3'and 5' of the cytosine dinucleotide were found to be critically important for A3G recognition. These observations provide the first demonstration that A3G cytosine deamination hotspots are defined by both sequence context and the single-stranded DNA secondary structure. This knowledge can be used to better trace the origins of mutations to A3G activity, and illuminate its impact on the generation of HIV-1 diversity, ultimately influencing the biological properties of the progeny virus variants.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF PUBLICATIONS

1. Clouser CL, **Holtz CM**, Mullett M, Crankshaw DL, Briggs JE, Chauhan J, Vanhoutan IM, Patterson SE, Mansky LM. Analysis of the ex vivo and in vivo antiretroviral activity of gemcitabine. PLoS One. 2011 Jan 14;6(1): e15840.

2. Clouser CL, **Holtz CM**, Mullett M, Crankshaw DL, Briggs JE, O'Sullivan MG, Patterson SE, Mansky LM. Activity of a novel combined antiretroviral therapy of gemcitabine and decitabine in a mouse model for HIV-1. Antimicrob Agents Chemother. 2012 Apr;56(4): 1942-8.

3. Dapp MJ, **Holtz CM**, Mansky LM. Concomitant lethal mutagenesis of human immunodeficiency virus type 1. J Mol Biol. 2012 Jun 8;419(3-4): 158-70.

4. **Holtz CM**, Mansky LM. Variation of HIV-1 Mutation Spectra Among Cell Types. J Virol. 2013 May;87(9):5296-9.

5. **Holtz CM**, Sadler H, Mansky LM. APOBEC3G cytosine deamination hotspots are defined by both sequence context and single-stranded DNA secondary structure. Nucleic Acids Research. 2013 Jul 1;41(12):6139-48

# CHAPTER I

# GENERAL INTRODUCTION

**History and Identification of Human Immunodeficiency Virus Type 1**

Human immunodeficiency virus type 1 (HIV-1) has killed over thirty million people and currently infects thirty-five million people worldwide (4). The virus infects and depletes cells of the immune system, which leads to a state of immunodeficiency referred to as AIDS (acquired immunodeficiency syndrome). AIDS is characterized by low CD4+ T-cell count and the presence of opportunistic infections such as oral thrush and toxoplasmosis or AIDS-related malignancies such as Kaposi's sarcoma or lymphoma. Once the disease progresses to AIDS, untreated individuals typically die within eleven months (128).

The first reported cases of AIDS (at the time called gay-related immune deficiency) were reported in the early 1980s when five gay men in the Los Angeles area were diagnosed with a rare lung infection PCP (*Pneumocystis carinii* pneumonia) (3). Around this same time an aggressive and rare form of cancer, Kaposi's sarcoma was reported in forty-one gay men in New York City and Los Angeles (2). At the time, the cause and origin of the opportunistic infections and cancer were not known. However, in 1983 HIV-1 was identified as the causative agent by Françoise Barré-Sinoussi, Luc Montagnier, and Robert Gallo (14, 79).

Further work by Robert Gallo, Françoise Barré-Sinoussi, and Luc Montagnier revealed that HIV-1 belonged to the *Retroviridae*, a family that includes other viruses such as HTLV (human T-cell leukemia virus), BLV (bovine leukemia virus), and MLV (murine leukemia virus) (14, 79). HIV-1 was further defined as belonging to the *lentivirus* genus of the *Retroviridae* (reviewed in (181)) which includes viruses such as visna virus, feline immunodefiency virus (FIV), and simian immunodeficiency virus

(SIV). Of these lentiviruses, SIV, a retrovirus of non-human primates, is believed to be the precursor to the virus we now know as HIV-1. Specifically, it is thought that HIV-1 was transferred from chimpanzees (i.e., $SIV_{cpz}$) to humans early in the 20$^{th}$ century in sub-Saharan Africa. An additional zoonotic event gave rise to a less transmissible form of HIV, referred to as HIV-2 and was likely transmitted to humans from sooty mangabeys (i.e., $SIV_{sm}$) in the mid 20$^{th}$ century.

**HIV-1 Transmission and Mechanism of Pathology**

Once transmitted from non-human primates to humans, HIV-1 was determined to be transmitted through sexual contact, breastfeeding, blood transfusions, contaminated needles, and perhaps through the oral cavity. However, transmission by sexual contact and breastfeeding requires HIV-1 to cross mucosal barriers in order to come into contact with target cells. To do this, HIV-1 "piggy-backs" on dendritic cells that transport the virus to target cells such as activated CD4+ T-cells (40, 88, 219). The dendritic cell is not infected by the founder virus, but instead brings HIV-1 into contact with CD4+ T-cells. The virus interacts with the receptors on the T-cell and infects it. Recently, HIV-1 was proposed to diffuse through squamous epithelium (37); however further studies are needed to confirm and extend these observations. In addition to infecting activated CD4+ T-cells, HIV-1 also infects resting CD4+ T-cells, macrophages, stromal fibroblasts, endothelial cells, gingival keratinocytes, astrocytes, fibroblasts in dental pulp and possibly dendritic cells (13, 30, 76, 81, 166, 197, 221).

Once HIV-1 has disseminated throughout the body, the virus depletes CD4+ T cells by at least four mechanisms including: 1) direct killing of infected cells; 2) decreased

naïve CD4+ cell production as a result of HIV-1 induced fibrosis of lymph tissue; 3) HIV-1-mediated damage to the lymph node fibroblastic reticular cell (FRC) network which decreases IL-7 production, thereby inducing apoptosis of naïve CD4+ cells (235), and; 4) apoptosis of CD4+ cells through chronic immune activation. Chronic immune activation refers to HIV-1-mediated depletion of gut-associated lymph tissue (GALT) (33, 86, 148). Loss of GALT enables bacterial constituents (i.e. lipopolysaccharide) to translocate the intestinal wall and gain entry into circulation (32). The presence of bacterial constituents in circulation causes systemic immune activation (32), which is characterized by release of proinflammatory cytokines and type I interferons. Type I interferon induces the expression of TRAIL (TNF-related apoptosis-inducing ligand) and its death receptor 5 in CD4+ T-cells in the periphery and secondary lymphatic tissues (100-102), which induces apoptosis of uninfected T-cells and destroys the lymph node architecture. Besides TRAIL, type I interferon also induces immunosuppressive enzymes (i.e., indoleamine-2,3-dioxygenase) that reduce T-cell proliferation and increase T-cell dysfunction (22, 23). The importance of chronic immune activation in the pathology of HIV-1 is demonstrated by the correlation between the degree of chronic immune activation and disease progression (80).

**HIV-1 Genome and Life Cycle**

To infect target cells and spread to other cells, as well as mediate immune destruction, HIV-1 relies on numerous viral proteins. Figure 1-1 depicts the HIV-1 genomic organization and Figure 1-2 illustrates a mature HIV-1 virion. Like all retroviruses, HIV-1 has three core genes including *gag*, *pol*, and *env*. HIV-1 *gag* encodes

for a polyprotein that has three structural domains: 1) capsid (CA) which forms the viral core, 2) nucleocapsid (NC) which encapsidates the viral RNA and protects it from degradation, and 3) matrix (MA) which surrounds and protects the viral core. The Gag polyprotein also includes smaller peptides, most notably p6, which plays important roles in the virus assembly pathway. The Gag polyprotein is cleaved into the mature protein products by the viral-encoded protease. The *pol* gene encodes a polyprotein that is translated from the full-length genomic RNA as a large Gag-Pol polyprotein that is created by ribosomal frameshifting. The Pol polyprotein encodes for three enzymes: 1) reverse transcriptase (RT), which is a RNA-dependent, DNA polymerase that converts ssRNA into dsDNA without proofreading activity 2) integrase, which integrates the viral dsDNA into the host genome, and 3) protease, which cleaves the Gag and Gag-Pol polyprotein, a necessary step in virus maturation. Finally, HIV-1, like all retroviruses, encodes for an envelope (*env*) gene that encodes for gp160, which is subsequently cleaved by a cellular protease into gp120 (a surface protein) and gp41(a transmembrane protein). The viral envelope protein (consisting of gp41 and gp120) is essential for binding the receptor on the target cell for infection. In addition to the conserved *gag*, *pol*, and *env* genes, HIV-1 encodes for six accessory proteins including Vpu (viral protein U), Nef (negative factor), Tat (transactivator of transcription), Vpr (viral protein R), Rev (regulator of viral protein expression), and Vif (viral infectivity factor). These accessory proteins facilitate viral replication and immune system evasion. Vpu induces the degradation of CD4 and counteracts the host protein tetherin (160, 222). Nef decreases the surface accumulation of MHC type I and II, which facilitates immune evasion. Tat allows for efficient elongation of transcription of the provirus. Vpr regulates the import

of the preintegration complex. Rev exports unspliced RNA from the nucleus. Finally, Vif counteracts the host protein APOBEC3G (apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like 3G) (203). Of the proteins encoded by the HIV-1 genome, the envelope protein is a key but not a sole determinant of cell tropism and initiates the first step in the viral life cycle (shown in Figure 1-3). Specifically, the envelope protein, gp120 interacts with CD4 receptors on the surface of host target cells. After binding to the CD4 receptor, gp120 undergoes a conformational change that exposes the fusion protein gp41, which interacts with the coreceptor CXCR4 or CCR5. The target cell and viral membranes fuse, and the viral core is deposited into the cytoplasm of the infected cell where the uncoating of the CA protein is initiated. During the uncoating of the viral core, the ssRNA genome is reverse transcribed into dsDNA by HIV-1 RT. The lack of proofreading activity makes the enzyme error prone which results in a high mutation rate of $3.4x10^{-5}$ mutations/target base pair/replication cycle (141), which is approximately one mutation per every three genomes. Viral dsDNA is then integrated into the host genome via the HIV-1 integrase to form the provirus. The viral genes are transcribed by cellular RNA polymerase II and translated by the host translational machinery. The full-length viral genomic RNA, the Gag and Gag-Pol polyproteins, and the Env proteins traffic to the plasma membrane to sites of virus assembly and release. After virus particles are released from the cell, the particle undergoes maturation due to the autocatalytic cleavage by viral protease and subsequent cleavage events to convert the immature particle into a mature and infectious particle with mature MA, CA and NC proteins as well as RT, protease and integrase. This maturation step gives rise to an infectious viral particle that can infect additional target cells. Additional infection events give rise to cells that are

actively infected and actively producing virus. However, a subset of infected cells harbor integrated proviruses but do not express viral RNA or proteins. These cells are called latently infected. At later points in time, some of these latently infected cells are reactivated (by mechanisms not completely understood), and active virus replication begins. Due to these latently infected cells, anti-HIV-1 therapy does not cure the infection from an infected individual.

**Drug Treatment to Inhibit HIV-1 Replication**

There are currently over 30 antiretroviral drugs approved by the Food and Drug Administration to treat HIV-1 (1). These drugs target steps within the HIV-1 life cycle including virus entry, reverse transcription, integration, and protease activity. Drugs that target viral entry do so by either blocking the HIV-1 coreceptor CCR5 (i.e., Maraviroc) or by preventing viral envelope gp41 fusion (i.e., Enfuvirtide). Maraviroc binds to CCR5 and interferes with the interaction of gp120 with CCR5. Enfuvirtide binds to gp41 and prevents the fusion of the viral and cell membranes. Anti-HIV-1 drugs that target reverse transcription fall into one of two classes including nucleoside/nucleotide-analog reverse transcriptase inhibitors (NRTIs) and non-nucleoside reverse transcriptase inhibitors (NNRTIs). Upon entry into cells, NRTIs are phosphorylated into their respective deoxyribonucleotide triphosphate (dNTP) analogs. The analog competes with endogenous dNTPs for incorporation into the viral DNA during reverse transcription. Once incorporated, NRTIs that lack a 3'-hydroxyl act as chain terminators to prevent elongation of viral DNA. In contrast, NNRTIs are allosteric inhibitors rather than competitive inhibitors. They bind to a hydrophobic pocket within reverse transcriptase,

which induces a conformational change that prevents elongation of the viral DNA (208). The exact mechanism by which reverse transcription is inhibited by NNRTI-induced conformational change is not known, however structural analysis of NNRTI-bound reverse transcriptase suggests that the conformational change may 1) change the position of the DNA primer in the active site (105), 2) decrease translocation of the primer/template and elongation of the nascent DNA strand (120, 215), or 3) prevent the positioning of dNTP relative to the primer/template (211, 231). These conformational changes disrupt enzyme structure by changing the geometry of the active site. The first integrase inhibitor (i.e., Raltegravir) interacts with the viral protein integrase, causing a delay in viral kinetics and subsequent inhibition. Finally, there are 11 approved protease inhibitors that target the viral protease and inhibit the Gag and Gag-Pol proteins from being processed. The virus particle cannot mature and therefore is unable to infect a new cell.

**HIV-1 Mutation Rate and the Emergence of Drug Resistance**

Current anti-HIV-1 drugs are given to patients to suppress viral replication. However, early on in HIV-1 treatment it was recognized that viral replication was only temporarily suppressed. For example, 3'-azidothymidine (AZT, a NRTI) only extended a patient's life about by six months when given as monotherapy. The lack of long-term efficacy was attributed to the rapid emergence of drug resistance. The speed at which the virus became drug resistant is in large part attributed to the high mutation rate of HIV-1 (along with the high rates of recombination and replication). HIV-1 has a high mutation rate of $3.4 \times 10^{-5}$ mutations/target base pair/replication cycle (141). Mutations can occur at

a few different steps of the life cycle including transcription of the genomic RNA by RNA polymerase II, viral DNA synthesis by reverse transcriptase, and proviral DNA replication by DNA polymerase delta (along with other cellular DNA polymerases). Unlike DNA polymerase delta, RNA polymerase II and reverse transcriptase do not possess proofreading ability. This lack of proofreading ability is thought to be the primary contributor to HIV-1 mutations acquired during replication, though the contribution of HIV-1 RT is thought to be significantly greater than that of RNA polymerase II (164). The high HIV-1 mutation, recombination, and replication rates lead to efficient immune evasion and drug resistance, and is the fundamental reason why no AIDS vaccine exists. As indicated above, in addition to the high mutation rate, HIV-1 replicates rapidly and it has been estimated that 1-10 billion virions are produced each day in an untreated individual (177). Given the high rates of replication and mutation, it is estimated that a mutation occurs at every nucleotide in the HIV-1 genome every day. In monotherapy, there is a high likelihood that a drug resistance-conferring mutation will develop rapidly considering the high mutation rate. By combining multiple drugs, there is less likelihood that a virus will emerge possessing all the mutations to confer resistance while maintaining viability. Therefore, HAART (highly active antiretroviral therapy) is a clinical drug regimen paradigm in which three or four drugs from different classes are administered to patients to lower the probability of the emergence of antiretroviral drug resistance. However, combination therapy decreases but does not eliminate the emergence of drug resistance. For example, 6-16% of newly infected patients have been infected with a virus resistant to at least one class of anti-HIV-1 drugs (189). HIV-1 develops resistance to every anti-retroviral by acquiring mutations that lessen or ablate

the drug's ability to inhibit the virus. Some mutations confer resistance to the specific drug being used or a few drugs, while a combination of mutations can cause resistance to all the drugs in the class (ie. NRTIs). The presence and prevalence of drug resistant HIV-1 emphasizes the need to continue to identify new HIV-1 drug targets.

**Mutation Rate as a Drug Target – Lethal Mutagenesis**

Studies with HIV-1 have demonstrated that virus variants with higher or lower mutation rates (i.e., mutator or antimutator phenotypes) have lower virus fitness (60). Such data help support the theory that HIV-1 infectivity can be reduced by altering mutation rate – particularly by the purposeful elevation of the HIV-1 mutation rate (26, 62). The HIV-1 mutation rate represents a potentially novel drug target that has yet to be fully exploited (50). While the high mutation rate of HIV-1 enables the virus to acquire drug resistance or evade the immune system, it also leads to a large proportion of non-infectious virions (See Figure 1-4) (52, 54, 59). An initial study showed that a 2-fold increase in picornavirus mutation frequency reduced viral titers by 99% (103). Subsequent studies showed that even a 3-4 fold increase in the mutation rate of HIV-1 and other RNA viruses renders the virus unable to replicate with enough fidelity to be viable or infectious, a processed referred to as lethal mutagenesis (54, 59, 69-71, 104, 132).

Lethal mutagenesis has been shown to be a plausible strategy for inhibiting viruses that have an inherently high mutation rate. For example, ribavirin has been shown to inhibit numerous RNA viruses *in vitro* (54, 123) through lethal mutagenesis. Ribavirin is incorporated as a nucleoside analog into poliovirus RNA increasing G-to-A and C-to-T

mutations (54, 55). Although ribavirin is used clinically to treat hepatitic C virus, it is not clear if the *in vivo* mechanism of action is lethal mutagenesis. There have been several compounds that have been shown to lethally mutagenize HIV-1. For example, Loeb et. al. first demonstrated lethal mutagenesis of HIV-1 using 5-hydroxydeoxycytidine (5-OH-dC ) (132). 5-OH-dC inhibits replication through lethal mutagenesis of HIV- 1 by inducing G-to-A mutations. However, 5-OH-dC is highly cytotoxic. Another mutagen, 5-aza-5,6,-dihydro-2'-deoxycytidine (KP-1212 or prodrug KP-1461) was shown to inhibit HIV-1 in cell culture and increase the frequency of G-to-A and A-to-G mutations (94). This was achieved with KP-1212 concentrations in the nanomolar to low micromolar range. KP-1461 advanced to clinical trials where the drug was shown to be safe *in vivo*. However, while KP-1461 did alter the mutation spectra *in vivo* there was not an increase in viral mutation rate *in vivo* and the clinical trials ceased. More recently, three FDA-approved anti-cancer drugs have been shown to increase the mutation rate of HIV-1; 5-azacytidine (5-AZC), gemcitabine and decitibine (50, 59). 5- AZC is a ribonucleoside analog of cytosine that is clinically used to treat myelodysplastic syndromes (107, 233). However, 5-AZC has also been shown to inhibit HIV-1 (27, 59) and this inhibition correlated with a threefold increase in mutation rate (59). Two mechanisms could account for the increase in HIV-1 mutation rate mediated by 5- AZC. For example, in its ribonucleoside form 5-AZC may be incorporated into RNA, which could cause a mispairing of the base during viral DNA synthesis and subsequent mutations (59). A second mechanism could be through incorporation of 5-AZC directly into the viral DNA. More specifically, 5-AZC could be either directly incorporated as a ribonucleotide or reduced into the 2'deoxynucleotide form by cellular ribonucleotide

reductase, which could then be incorporated into viral DNA in place of dCTP. After incorporation into viral (–) sense DNA, the unstable 5-AZC triazine ring has been hypothesized to open due to a disruption of a bond between nitrogen-1 and carbon-6 and subsequent hydrolysis. The ring opening allows base pairing with a cytosine that results in a G-to-C mutation. Dapp et. al. observed a higher mutant frequency during the early phase of replication attributing to the reverse transcription step of the viral life cycle (59). Therefore, it was concluded that 5-AZC causes lethal mutagenesis by incorporation into the viral DNA.

Another nucleoside analog that causes lethal mutagenesis *in vitro* is decitabine (i.e., 5-aza-2'-deoxycytidine) (50). Like 5-AZC, treatment of HIV-infected cell with decitabine led to the observation of G-to-C transversion mutations in HIV-1 proviruses. Decitabine is the reduced form of 5-AZC and is thought to be directly incorporated into viral DNA during reverse transcription. *In* vivo, decitabine has been shown to inhibit DNA methylation and is approved clinically for the treatment of myelodysplastic syndrome (151-153). Decitabine has been shown more recently to reduce HIV-1 infectivity (50). The anti-HIV-1 activity of decitibine was shown to correlate with an increase in the HIV-1 mutation rate, specifically causing an increase in G-to-C transversion mutations in the viral DNA. Due to the ability to be incorporated into DNA, mutagens are thought to have high cellular toxicity and carcinogenicity. More specifically, decitabine has been shown to cause G-to-C mutations in the colonic DNA of transgenic mice containing *Escherichia coli lac I* transgene (109). However, these mutations have not been observed in host genomes when used clinically, and decitabine was shown to not be genotoxic (38, 50, 153).

Besides mutagenic nucleoside analogs, nucleoside analogs that alter cellular dNTP pools have also been shown to increase HIV-1 mutation rate. One of these nucleoside analogs, gemcitabine, alters dNTPs by inhibiting host ribonucleotide reductase. Ribonucleotide reductase is the enzyme that converts rNTPs into dNTPs. If this enzyme is inhibited, then dNTP pools in the cell are imbalanced and the cell may be unable to complete DNA replication or increase the mutation rate by incorporation of incorrect bases. Gemcitabine is a clinically approved drug for cancer therapy that is a ribonucleotide reductase inhibitor. Gemcitabine has been shown to decrease infectivity of HIV-1 at drug concentrations without apparent toxicity in a cell culture system (50). Interestingly, when gemcitabine was combined with the mutagen decitabine at low concentration, there was a synergistic reduction of HIV-1 infectivity (50) that correlated with an increase in the mutant frequency. A synergistic decrease in HIV-1 replication was also seen when decitabine was combined with another ribonucleotide reductase inhibitor, hydroxyurea (50). However, hydroxyurea has been shown to be too toxic to be used clinically to treat HIV-1. The ability to use decitabine in combination with other ribonucleotide reductase inhibitors such as gemcitabine enables each drug to be used at lower concentrations than would be required if the drugs were used individually and thereby potentially reduces the toxicity to the cells.

**Lethal Mutagenesis of HIV-1 by APOBEC Host Proteins**

Some cellular proteins have evolved to restrict retroviral replication by inducing lethal mutagenesis. For example, the APOBEC (apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like) family of proteins has been shown to lethally

mutagenize HIV-1, as well as induce mutations in both ssDNA and ssRNA. The APOBEC family are cytidine deaminases that include 11 members; AID (activation-induced deaminase), APOBEC1, APOBEC2, APOBEC3A-D/F-H, and APOBEC4. Of these, AID induces somatic hypermutation and class switch recombination by deaminating the switch region in the ssDNA of B-cells within the germinal centers (31, 65, 134). In contrast, APOBEC1 deaminates a cytosine in apolipoprotein B mRNA, which creates a truncated form of the mRNA that is important for the absorption and transport of dietary fats (228), while the function of APOBEC2 and APOBEC4 is unknown. The APOBEC3 proteins have both anti-retroelement and antiviral activity. Some APOBEC3 members have been shown to inhibit both LTR (long terminal repeat) retrotransposons and non-LTR retrotransposons including long interspersed nuclear elements (LINEs) (24, 25, 46, 48, 75, 117, 156, 200, 213).

Structurally, there is high amino acid identity between the APOBEC3 family members with as much as 50% sequence homology. The APOBEC3 family has two different protein structures with APOBEC3A, APOBEC3C, and APOBEC3H containing a single catalytic domain, while APOBEC3B, APOBEC3D, APOBEC3F, and APOBEC3G have two catalytic domains. The catalytic domains are zinc coordinated and highly conserved between the APOBEC3 family containing a His-X-Glu-$X_{23-28}$-Pro-Cys-$X_{2-4}$-Cys cytosine deaminase motif (110, 228).

Evolutionarily, the APOBEC3 proteins expanded in humans to seven compared to mice where there is only one APOBEC3 protein (110). It is hypothesized that in primates the gene duplicated several times creating the expanded locus. Interestingly, the expansion of the APOBEC3 locus coincides with the decreased activity of retroelements

in primates. This decrease in retroelement activity may explain the finding that while humans have five times more retroelements in their genome than mice, their retroelement activity is 100 times lower (137). This suggests that APOBEC3 proteins may be important in regulating retroelement activity. Cell type expression of APOBEC3 proteins vary depending on the specific APOBEC. For example, APOBEC3A is highly expressed in monocytes/macrophages while APOBEC3B is expressed in keratinocytes and the small intestine. APOBEC3C, APOBEC3D, APOBEC3F, and APOBEC3G are all expressed in CD4+ T-cells while APOBEC3H is expressed in peripheral blood lymphocytes and the testis/ovaries. Expression of the APOBEC proteins within their respective cell types may restrict and/or contribute to the evolution of viruses.

**APOBEC-Mediated Restriction of HIV-1 Replication**

The APOBEC3 proteins have been implicated in restricting numerous viruses including adeno-associated virus, hepatitis B virus, human papilloma virus, murine leukemia virus, and most notably, HIV-1 (21, 46, 68, 95, 118, 140, 202, 214, 218, 224). In fact, HIV-1 is restricted by APOBEC3D, APOBEC3F, APOBEC3G, APOBEC3H, and modestly by APOBEC3C and the inhibition of HIV-1 is attributed to the catalytic activity (58, 106, 124, 240). The most studied APOBEC3 proteins, APOBEC3F and APOBEC3G, are packaged into the viral core during assembly through possible interactions with the structural protein NC and RNA (6, 41, 198, 237). HIV-1 has developed a mechanism to counteract an APOBEC3 assault through its accessory protein Vif. HIV-1 Vif interacts with APOBEC3 proteins and targets them for proteosomal degradation by the E3 ligase Ubiquination complex (203). However, in the absence of

Vif, APOBEC3F and APOBEC3G deaminate cytosine residues to uracil residues in the minus (–) sense DNA during reverse transcription. These uracil residues in the viral DNA attract the enzymes uracil DNA glycosylase and apurinic-apyrimidinic endonuclease, degrading the viral DNA. As shown in Figure 1-5, the viral minus (-) sense DNA that escapes degradation completes reverse transcription resulting in the uracil residue base pairing with an adenine residue causing G-to-A mutations in the plus (+) sense DNA (95, 234). As a result, the high amount of mutations, termed hypermutation, may alter open reading frames and introduce translational stop codons ceasing replication. APOBEC3 proteins preferentially mutate cytosine residues within a given sequence context. For example, APOBEC3G targets C<u>C</u> motifs in the (-) strand of DNA or what is observed as a GG dinucleotide in the (+) strand of DNA (underlined letter indicates mutation target). The other members of the APOBEC3 family recognize T<u>C</u> motifs or GA dinucleotides. Interestingly, not all CC or TC dinucleotides are mutated by the APOBEC3 proteins. The reasons why some dinucleotides are targeted and others are not is largely unknown.

Besides their ability to mutate viruses via cytosine deaminase activity, APOBEC3 proteins have also been reported to restrict viruses through non-catalytic activity. These functions include physical impairment of reverse transcription, decreasing the ability of $tRNA^{Lys3}$ primers to initiate reverse transcription, and reducing $tRNA^{Lys3}$ cleavage during the plus strand DNA transfer step of reverse transcription (20, 87, 147, 161, 162). However, these non-catalytic functions were discovered in an overexpression system where the APOBEC3 was not as physiological relevant levels leading some to doubt the *in vivo* relevance.

Although APOBECs have been shown to lethally mutagenize HIV-1, APOBEC3G has also been shown to sublethally mutate HIV-1 (192). Sublethal mutations occur at low A3G proteins levels. These sublethal mutations may enhance the ability of the virus to survive selective pressure and facilitate the emergence of drug resistance. As HIV-1 drug resistance becomes increasingly prevalent, novel approaches and drugs are needed to treat virus infection. Lethal mutagenesis is a novel strategy to inhibit HIV-1 replication since none of the current anti-HIV-1 drugs inhibit replication by this mechanism. A better understanding of the mechanism by which APOBECs inhibit HIV- 1 replication or contribute to viral evolution is important for understanding how to aggressively combat the virus and inhibit its replication.

**Dissertation Objectives:**

The goals of the following studies were to understand the effect of cell type, mutagenic drugs and host factors on HIV-1 evolution and infectivity. Specifically, Chapter 2 examines if HIV-1 variation is cell type dependent while Chapter 3 sought to elucidate the nucleotide-specific context that is required for APOBEC3G-induced mutations. The results presented in Appendices I and II examine the efficacy and toxicity of the ribonucleotide reductase inhibitor, gemcitabine, alone and in combination with the mutagen, decitabine, in a murine AIDS model. Finally, Appendix III addresses how two mutagens, APOBEC3G and 5-AZC, that both target cytosine residues interact to decrease HIV-1 infectivity through lethal mutagenesis.

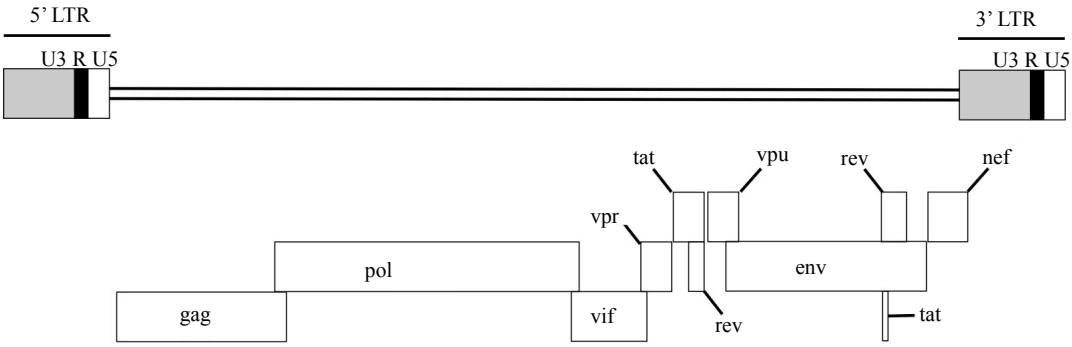



**Figure 1-1. HIV-1 genome.** Diagram of the human immunodeficiency virus type 1 (HIV-1) proviral DNA and gene organization. The 5' and 3' long terminal repeats (LTR) are located at the ends of the HIV-1 proviral sequence and contain transcriptional regulatory elements within unique 3' region (U3), repeat region (R), and unique 5' region (U5). The HIV-1 proviral DNA encodes for structural (*gag*), enzymatic (*pol*), envelope (*env*), and accessory genes (*vif, vpr, tat, vpu, rev, nef, tat*). The Gag protein contains matrix, capsid and nucleocapsid domains (not shown). The *pol* gene encodes for the viral protease, reverse transcriptase and integrase (not shown).

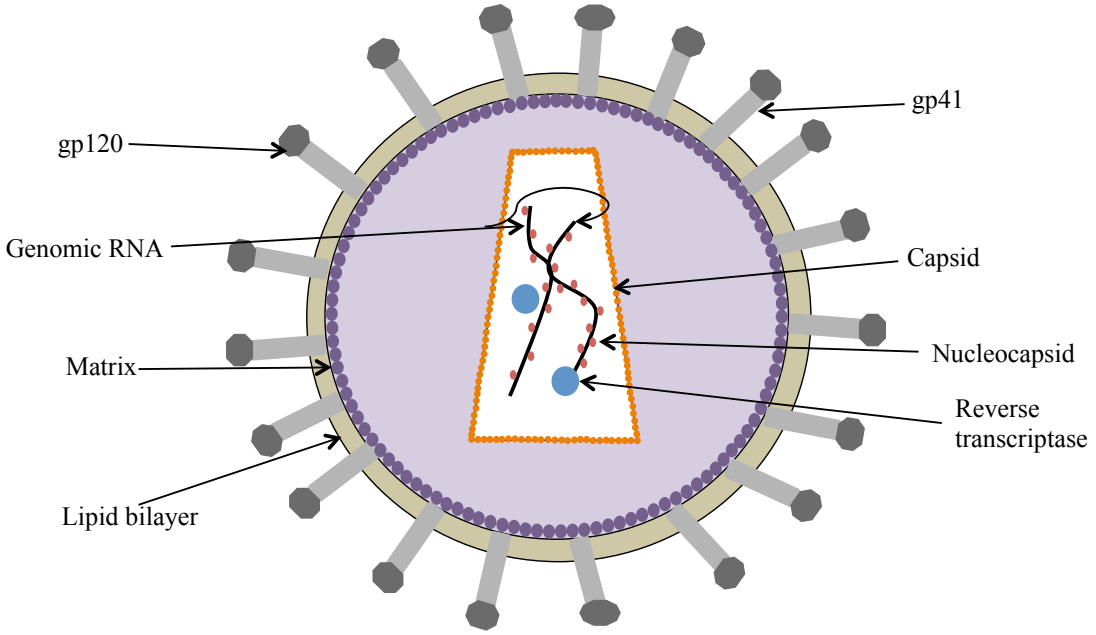



**Figure 1-2. Mature HIV-1 particle.** Shown is a cartoon depiction of a cross section through a HIV-1 viral particle. The envelope proteins (gp120 and gp41) are shown as a dark grey circle or a grey rectangle respectively. Matrix protein is shown as dark purple circles that lines the inside of the lipid bilayer which is in grey-yellow. The capsid protein is shown as small orange circles and forms the viral core. The two copies of genomic RNA, which are depicted as black lines, are contained within the CA and are coated by nuclecapsid protein, shown as red circles. Reverse transcriptase is also located within the viral core and shown as as blue circles.

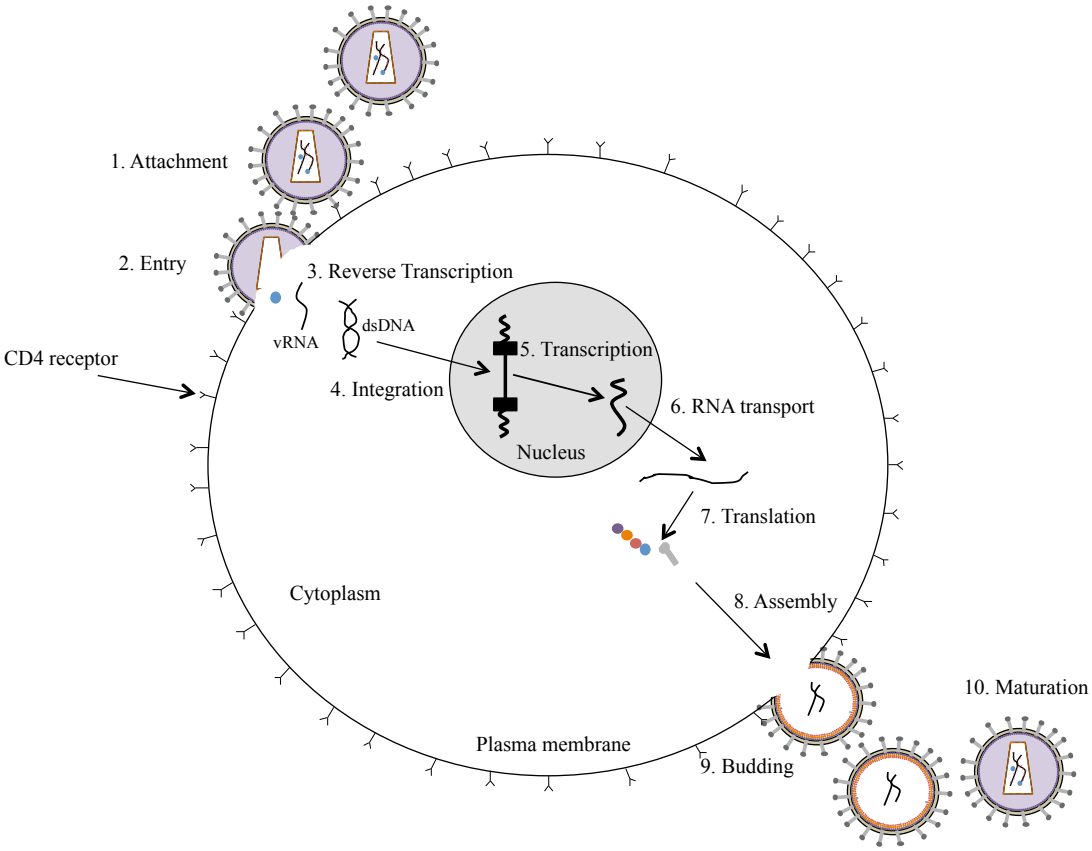



**Figure 1-3. HIV-1 life cycle.** A mature HIV-1 viral particle interacts with and attaches to a CD4 receptor, and a CXCR4/CCR5 coreceptor (not shown), via the envelope protein gp120 (step 1). The virion fuses with the host cell through gp41 interactions and deposits its core with the RNA genome into the cell (step 2). The viral RNA (vRNA) genome undergoes reverse transcription via reverse transcriptase (step 3) and is converted into dsDNA that is transported through the cytoplasm and then imported into the nucleus where it is integrated into the host genome via integrase (step 4). The HIV-1 proviral DNA is transcribed by RNA polymerase II (step 5) and transported out of the nucleus (step 6). The viral mRNA is translated by host cell translation machinery to produce the viral proteins (step 7). Shown are the translation of the structural protein (Gag), reverse

transcriptase and envelope (gp120 and gp41). The viral proteins assemble at the membrane of the cell (step 8). The immature virion buds (step 9) and undergoes a maturation process via viral protease (step 10) forming the viral core and rendering the viral particle infectious.

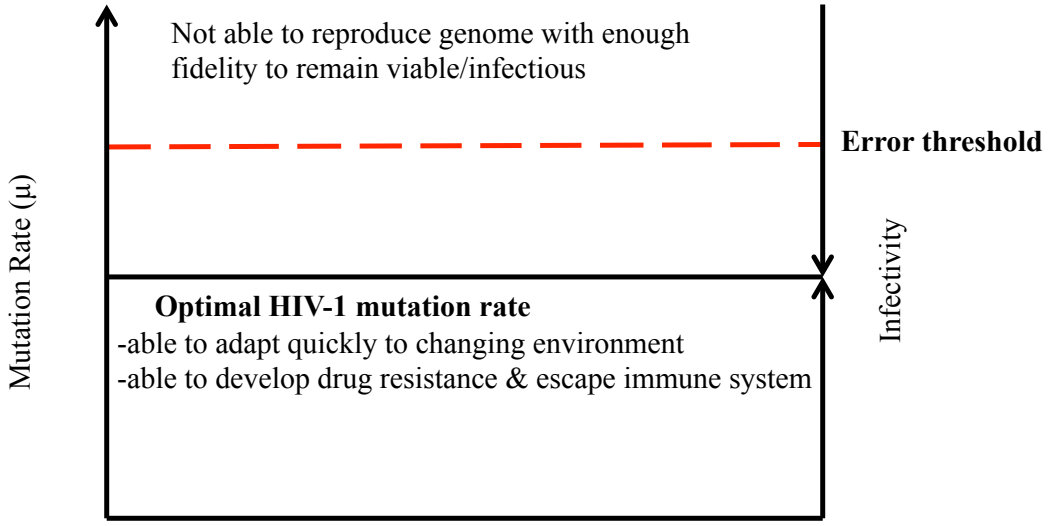



**Figure 1-4. HIV-1 lethal mutagenesis.** Shown is a figure describing lethal mutagenesis. HIV-1 has an optimal mutation rate in a given environment (shown as a black solid line) where the virus adapts to the changing environment and is able to escape the immune system. The left y-axis shows the mutation rate (μ) and its relationship to HIV-1 infectivity on the right y-axis. As the mutation rate decrease, the infectivity decreases due to a lack of ability to adapt quickly to the environment. As the mutation rate increases, the infectivity also decreases due to the additional acquisition of mutations. Above a certain mutation rate, known as the error threshold (shown as a dotted red line),

the viral genome acquires too many mutations rendering the virus unviable (ie. lethal mutagenesis). Adapted from Dapp *et al*, 2013 and C. Clouser (unpublished).

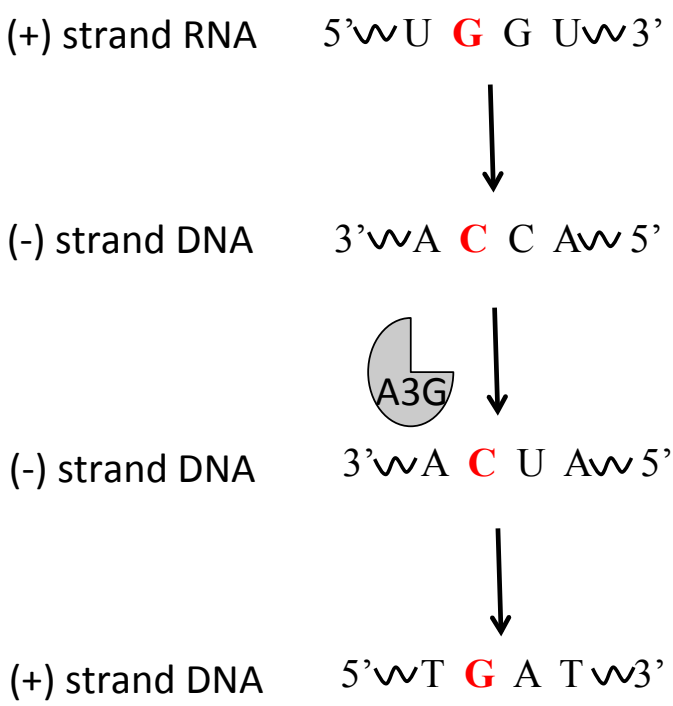



**Figure 1-5. APOBEC3G-mediated mutagenesis of HIV-1 DNA.** The plus (+) strand RNA genome is converted into minus (-) strand DNA by reverse transcriptase. The CC dinucleotide in the (-) strand DNA is recognized by APOBEC3G and the first cytosine is deaminated to a uracil (U). The uracil base pairs with an adenine (A) in the (+) strand DNA; this results in a G-to-A transition mutation in the (+) strand DNA. The "red" highlighted nucleotide base is the site of APOBEC3G –mediated mutagenesis.

# CHAPTER II

# VARIATION OF HIV-1 MUTATION SPECTRA AMONG CELL TYPES

**Text**

Human immunodeficiency virus type 1 (HIV-1) infection remains a serious threat to global public health, with over 34 million people infected worldwide (http://www.unaids.org). In the absence of an effective vaccine, antiretroviral drug therapy remains the primary means for preventing transmission and disease progression, as well as new infections (92). The error-prone nature of HIV-1 reverse transcriptase (RT) remains a key determinant in the generation of mutations during HIV-1 replication. HIV-1 RT plays a prominent role in the high genetic diversity and evolution of retroviruses, which are driven by high rates of viral replication, mutation, and recombination (142). High genetic diversity allows HIV-1 to evade the immune system and become resistant to drug therapy. This high mutation rate and the high rate of virus evolution impact virus transmission and disease progression, as well as promote the emergence of antiviral drug resistance (142).

Cellular proteins can exploit the propensity of HIV-1 to rapidly mutate by enhancing viral mutagenesis to the point where the virus is unable to replicate with enough fidelity to remain infectious (62). The apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like 3 (APOBEC3) family of proteins, for instance, act as cytidine deaminases during reverse transcription, causing G-to-A hypermutation in Vif-deficient HIV-1 (202, 240). APOBEC3G (A3G) has been particularly well characterized for its ability to induce cytosine deamination during the minus-strand DNA synthesis step of reverse transcription, resulting in G-to-A mutations in the plus-strand viral DNA (95, 234). The editing activities of APOBEC3G result in G-to-A hypermutation (and lethal mutagenesis), although sublethal mutagenesis has been demonstrated (192).

APOBEC3D, APOBEC3F, and APOBEC3H have also been demonstrated to restrict and cause G-to-A hypermutation of Vif-deficient HIV-1 (106).

Previous studies have provided evidence that HIV-1 genetic variation is impacted by RNA polymerase II transcription errors as well as minus-strand and plus-strand mutations that arise during HIV-1 reverse transcription (143, 158, 239). Deoxynucleoside triphosphate (dNTP) levels have been shown to have direct effects on RT fidelity (17, 113). Macrophages have been shown to have low dNTP levels, and this decreases the efficiency of viral DNA synthesis and increases the likelihood of mutations occurring during HIV-1 reverse transcription (66). A previous study investigated the differences in the HIV-1 mutation rate and mutation spectra observed between HIV-1 replication in HeLa cells versus that of CEM-A cells (141). No significant differences were observed in the viral mutation rates between these two cells. While the number of mutants that were characterized was small and prohibited a statistical analysis of potential differences in the rates of mutation for specific mutation types, the general trends suggested that there was no significant difference in the rates of base pair substitution mutations, frameshift mutations, and deletion or deletion-with-insertion mutations (141).

In the present study, we sought to further investigate the role of how cell type affects the mutation rate of HIV-1 as well as the virus mutation spectra. For analysis of mutation rate and mutation spectra, CEM-GFP (green fluorescent protein-expressing T lymphoblast cell line; NIH AIDS Research and Reference Reagent Program via J. Corbeil), U373-MAGI-CXCR4$_{CEM}$ cells (glioblastoma cell line; via M. Emerman through the NIH AIDS Research and Reference Reagent Program), 293T cells (human embryonic kidney cells; American Type Culture Collection [ATCC]) and SupT1 cells (T

lymphoblast cell line; ATCC) were transduced with an HIV-1 vector, pHIG (192). Mutant detection was determined by detection of cells that had a mouse heat-stable antigen-negative ($HSA^-$) $GFP^+$ phenotype. The calculation of mutant frequency was determined by dividing the number of infected cells harboring a mutated provirus phenotype (i.e., $HSA^-$ $GFP^+$) by the total number of cells infected (i.e., $HSA^-$ $GFP^+$ and $HSA^+$ $GFP^+$). Using this analysis, it was determined that there was no significant difference ($P > 0.05$, Student's t test) in the frequencies of mutants recovered from these cell types in parallel analyses (Table 2-1). A similar lack of observed difference in mutation rates had previously been made in an analysis conducted comparing HeLa cells and CEM-A cells (141).

To determine if the HIV-1 mutation spectra were influenced by cell type, proviral HSA mutation target sequences were analyzed from the four cell lines after parallel virus infections. Interestingly, a significant difference was noted when comparing T-to-C transition mutations between CEM-GFP and SupT1 cells ($P < 0.05$, Fisher's exact test) and G-to-A transition mutations between CEM-GFP and 293T ($P < 0.01$), CEM-GFP and SupT1 ($P < 0.01$), and U373-MAGI and 293T ($P < 0.05$) (Table 2-2) cells. No other significant differences ($P > 0.05$) between any of the other mutation types were detected in these analyses. The relatively high level of A-to-G mutations observed in all 4 cell lines could be due to either ADAR-1 or ADAR-2 activity (72, 73, 180). When sequences that possessed multiple G-to-A mutations in the proviral HSA mutation target gene sequence were removed from the analysis, no significant differences ($P > 0.05$) in G-to-A mutations among the 4 cell types analyzed were observed (data not shown). This indicates that the frequency of G-to-A transition mutations was significantly influenced

by the proviral sequences recovered that contained multiple G-to-A mutations in the target gene sequence.

Proviral sequences with multiple G-to-A mutations in the target gene sequence were recovered from CEM-GFP, U373-MAGI, and 293T cells (Table 2-3). There were significant differences (Fisher's exact test) in the frequencies of sequences with multiple G-to-A mutations recovered between CEM-GFP and 293T cells (P = 0.03), CEM-GFP and SupT1 cells (P = 0.0005), and U373-MAGI and SupT1 cells (P = 0.006). The mutational load for G-to-A mutations per proviral sequence from each of the cell lines was analyzed (Fig. 2-1A). Although the average numbers of G-to-A mutations per sequence were not statistically different between CEM-GFP, U373-MAGI, and 293T cells, there was a trend toward a higher degree of G-to-A mutations per sequence in CEM-GFP and U373-MAGI cells relative to that of 293T cells. The locations of G-to-A mutations analyzed indicated that certain G residues were hot spots for G-to-A mutations (Fig. 2-1B). Of these mutations, 73% of the G-to-A mutations occurred at GA dinucleotides, while 17% occurred at GG dinucleotides, 5% occurred at GT dinucleotides, and 5% occurred at GC dinucleotides.

The observation that G-to-A mutations preferentially occurred at GA dinucleotides suggested that these mutations may be due to the expression of APOBEC3 proteins. In order to test whether there was a correlation between the frequency of recovered proviral sequences with multiple G-to-A mutations and that of APOBEC3 gene expression, mRNA expression levels of the APOBEC3 proteins were analyzed in each of the cell lines under study. The study of mRNA expression levels was done as a surrogate for assessment of APOBEC3 protein levels, given that antibodies that can readily

differentiate the family of proteins are not currently available. A3A, A3B, A3C, A3D/E, A3F, A3G, A3H, and TATA-binding protein expression vectors were used as control standards (184). Quantitative reverse transcriptase PCR (qRT-PCR) analysis was done on each cell line for A3A, A3B, A3C, A3D/E, A3F, A3G, and A3H (Fig. 2-1C). A3C mRNA expression levels suggested a possible correlation with the prevalence of proviral sequences harboring multiple G-to-A mutations from each cell line (i.e., CEM-GFP > U373-MAGI > 293T > SupT1). However, analysis using a Pearson correlation coefficient test did not support this correlation ($P = 0.052$). The A3G mRNA expression also appeared to have an expression pattern that appeared to correlate to the proviral sequences with multiple G-to-A mutations from each cell line. In particular, significant differences were observed in A3C expression between CEM-GFP and SupT1 ($P = 0.0015$, Student's t test), CEM-GFP and 293T ($P = 0.0213$), U373-MAGI and 293T ($P = 0.0286$), U373-MAGI and SupT1 ($P < 0.001$), and 293T and SupT1 ($P < 0.001$) cells. However, the observed dinucleotide specificity (i.e., GA) was not consistent with the dinucleotide specificity of A3G (i.e., GG). All of the other APOBEC3s either lacked mRNA expression in one of the cell lines where multiple G-to-A mutations were observed, or there was no significant difference in mRNA levels for a particular A3 between the 4 cell lines analyzed. It is interesting to note that while both CEM-GFP and Sup T1 cells are T lymphoblast cell lines, these findings indicate that distinct virus mutation spectra were observed. Taken together, the origins of the multiple G-to-A mutations in recovered proviral sequences could not be directly attributed to A3 gene expression in the 4 target cell types.

To date, the nature of how cell type influences the mutation rate and mutation

spectra of HIV-1 has not been extensively studied. The discovery of the APOBEC3 proteins provided clear evidence of the potential for the host cell to extensively edit and mutate the HIV-1 genome, which may lead to mutations that can shape HIV-1 evolution (192). In this study, we have investigated several cell lines that are commonly used to study HIV-1 replication in cell culture. Using an HIV-1 vector to study the mutation rate and spectra, we observed that cell type did not influence the viral mutation rate, similar to an analysis of HIV-1 mutation rates between HeLa and CEM-A cells (141). Importantly, we observed for the first time distinct differences in HIV-1 mutation spectra in parallel analyses of 4 different target cell types. In particular, there was a significant difference in the frequency of G-to-A transition mutations observed between CEM-GFP and 293T cells, U373-MAGI and 293T cells, and between CEM-GFP and SupT1 cells. A significant difference in T-to-C transition mutations was also observed between CEM-GFP and SupT1. Interestingly, analysis of the proviral sequences from CEM-GFP and U373-MAGI cells led to the characterization of sequences harboring multiple G-to-A mutations in the reporter gene sequence, ranging from 2 to 13 G-to-A mutations. Furthermore, there were significant differences in frequencies between CEM-GFP and 293T cells, CEM-GFP and SupT1 cells, and U373-MAGI and SupT1 cells. The proviral sequences with multiple G-to-A mutations occurred mainly at GA dinucleotides (i.e., 73%) and GG dinucleotides (i.e., 17%).

Analysis of APOBEC3 mRNA expression levels did not allow for a correlation between the expression level of any one particular APOBEC3 and the observed sequences with multiple G-to-A mutations. There have been previous reports implicating an APOBEC3 protein as having a target cell effect in generating G-to-A mutations (28,

121). Koning et al. hypothesized that A3A was responsible for G-to-A editing of HIV-1 cDNA in macrophages (121), while Bourara et al. hypothesized that A3C could mutate HIV-1 viral DNA in the target cell (28). In both studies, the G-to-A mutations were likely sublethal and did not result in G-to-A hypermutation. It is possible that in these previous studies, G-to-A hypermutation did occur but was at a low frequency and was not detected.

Our studies here are distinct in that mutation spectrum differences involving both sequences containing multiple G-to-A mutations in the mutation target gene sequence (which is indicative of G-to-A hypermutation), and differences in the frequencies of other mutation types were observed. While the origins of these mutations were not determined, differences in cellular protein levels that could edit HIV-1 DNA or differences in the fidelity of HIV-1 reverse transcriptase and/or cellular RNA polymerase II are likely responsible for the changes in mutation spectra observed in this study. As indicated earlier, differences in nucleotide pool levels in various cell types can influence the likelihood of RT-mediated mutation (66). Further studies to determine the molecular basis for these differences in mutation spectra would be of particular interest, as well as studies that would extend these studies to human primary cells (e.g., primary T cells and macrophages). Additional studies to investigate the ability of APOBEC3 proteins are warranted and would be enhanced with antibody reagents that allow for specific detection of each APOBEC3 protein as well as specific small interfering RNAs (siRNAs) for mRNA depletion studies.

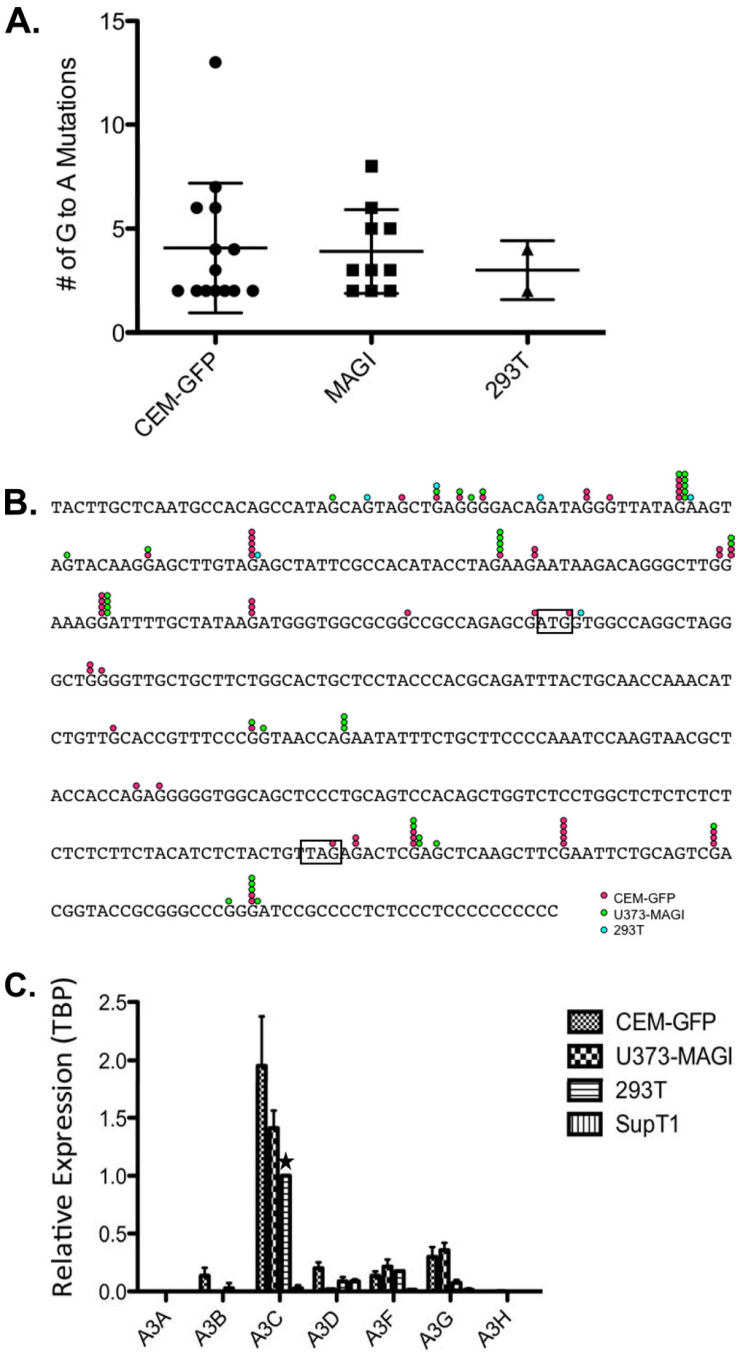



**Figure 2-1. G-to-A mutational load and mutation location in HIV-1 proviruses with multiple G-to-A mutations.** (A) G-to-A mutational load in HIV-1 proviruses from CEM-GFP, U373-MAGI, and 293T cells. Each provirus with a mutant HSA sequence

containing with multiple G-to-A mutations is indicated by a black circle (CEM-GFP), black square (MAGI), or black triangle (293T). The average G-to-A mutational load and standard deviation are indicated. (B) Location of G-to-A mutations in recovered proviruses harboring multiple G-to-A mutations. The red, green, and blue circles above G residues indicate the locations of G-to-A mutations (one circle per mutation identified) in proviruses recovered from CEM-GFP, U373-MAGI, and 293T, respectively. The start and stop codons of the HSA gene are identified by black rectangular boxes. (C) Quantitative RT-PCR was performed to determine the relative levels of APOBEC3 mRNA expression among the cell lines under investigation (i.e., CEM-GFP, SupT1, U373-MAGI, and 293T). The asterisk indicates that the APOBEC3C level from 293T cells was set to 1, and all other values are relative to this measurement. The mRNA expression levels were normalized to TATA-binding protein (TBP) mRNA levels. For these analyses, data with a difference in efficiencies between the APOBEC3 standard and TBP standard were $\leq$10% and $R^2 \geq 0.98$. Threshold cycle (CT) values were determined using the regression method, and data were analyzed by $E^{-\Delta CT}$. Experiments were conducted in triplicate with the standard deviation indicated.

**Table 2-1.** Mutant frequency of HIV-1 among selected cell lines[a]

| Cell Line | Mutant Frequency |
|---|---|
| CEM-GFP | 0.11 ± 0.04 |
| U373-MAGI | 0.12 ± 0.03 |
| 293T | 0.11 ± 0.04 |
| SupT1 | 0.11 ± 0.01 |

[a] Two million 293T cells were transfected with 10ug of pHIG and 1ug of pVSV-G using the calcium phosphate method as previously described (74). The viral supernatants were collected 48h posttransfection, filtered, and used to infect permissive target cells ($1.4\times 10^6$). Twenty-four hours postinfection, the cells were harvested and prepared for flow cytometry as described previously (192). Infected cells, typically 20 to 40% of the total for each target cell line, were analyzed by flow cytometry for expression of two marker genes, coding for HSA and GFP. Mutant frequencies were calculated from the cell population phenotypes identified by flow cytometry using the formula $(HSA^- GFP^+)/[(HSA^- GFP^+) + (HSA^- GFP^+)]$. For each target cell line, ~$7\times10^5$ to $9\times10^5$ cells were identified as infected cells (i.e., $HSA^- GFP^+$ and $HSA^+ GFP^+$ cell populations) and ~$8\times10^4$ cells were identified as infected cells with a mutation in the reporter gene (i.e., $HSA^- GFP^+$).

**Table 2-2.** Mutation spectra in the HSA mutation target gene of HIV-1 proviral sequences[a]

| | CEM-GFP[a] | U373-MAGI[b] | 293T[c] | SupT1[d] |
|---|---|---|---|---|
| **G-to-A** | 29% | 26% | 17% | 17% |
| **A-to-G** | 27% | 29% | 29% | 29% |
| **T-to-C** | 27% | 29% | 30% | 36% |
| **C-to-T** | 10% | 10% | 14% | 11% |

[a] Cells were transduced with an HIV-1 vector (HIG) pseudotyped with vesicular stomatitis virus protein G (VSVG), and the HSA reporter gene sequence from the provirus in infected cells was PCR amplified and sequenced. The percentage of mutations for each mutation type compared to the total mutations identified is indicated.
[b] Total no. of sequences, 238; total no. of mutations, 365.
[c] Total no. of sequences, 225; total no. of mutations, 329.
[d] Total no. of sequences, 154; total no. of mutations, 199.
[e] Total no. of sequences, 173; total no. of mutations, 219.

**Table 2-3.** Proportion of HIV-1 proviral sequences recovered per infected cell line analyzed possessing multiple G-to-A mutations in the HSA mutation target gene[a]

| Target Cells | No. (%) of mutant sequences with multiple G-to-A mutations/total mutants analyzed |
|---|---|
| CEM-GFP | 14/238= 5.9% |
| U373-MAGI | 10/225= 4.4% |
| 293T | 2/155= 1.3% |
| SupT1 | 0/173= 0% |

[a] Cells were transduced with an HIV-1 vector (HIG) pseudotyped with vesicular stomatitis virus protein G (VSVG), and the HSA reporter gene sequence from the provirus in infected cells was PCR amplified and sequenced. The number of proviral sequences having HSA reporter gene sequences with multiple G-to-A mutations from each cell type was divided by the total number of proviral sequences identified with mutations.

# CHAPTER III

# APOBEC3G CYTOSINE DEAMINATION HOTSPOTS ARE DEFINED BY BOTH SEQUENCE CONTEXT AND SINGLE-STRANDED DNA SECONDARY STRUCTURE

## Introduction

Apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like 3G (APOBEC3G or A3G) is an important host restriction factor that can inhibit human immunodeficiency virus type 1 (HIV-1) and other viruses via cytosine deamination of viral genomic DNA (138). In the presence of the HIV-1 vif protein, the activity of APOBEC3G is attenuated and the residual deamination activity of A3G may contribute to the high mutation rate of HIV-1, virus evolution and antiretroviral drug resistance (90, 95, 111, 140, 192). However, when the activity of vif is moderated or extinguished, A3G highly restricts viral replication (83, 96, 138). This restriction largely results from the high level of deamination during HIV-1 reverse transcription, which can lead to degradation of DNA with abasic residues, a decrease in the specificity of plus-strand initiation, and accumulation of lethal G-to-A mutations on the plus strand (i.e., hypermutation) (95, 140, 238). Although it is known that A3G acts exclusively on single-stranded DNA (ssDNA) and acts preferentially at specific sites in a sequence of DNA, termed 'hotspots,' the factors that create these critical restriction hotspots are not fully understood (119, 182). A3G requires a cytosine dinucleotide context on ssDNA and studies have shown that A3G tends to favor CCCA or T/CCC sequences (21, 234). Distinct restriction hotspots on the viral genome often occur outside this four base context, however, as A3G can deaminate at a variety of other sites (115). Interestingly, many CCCA or T/CCC sites are not edited by A3G (192).

As indicated above, published observations to date indicate that hotspot specificity for A3G must be determined by more than the bases immediately 3′ and 5′ of the required cytosine dinucleotide sequence. Specific sequences far upstream or

downstream from a hotspot also cannot be necessary for recognition, since A3G can deaminate oligos as small as 16 or 13 nt efficiently *in vitro* (108, 216). However, it is formally possible that certain distal sequences could play some role in large ssDNAs *in vivo*. Given that cytosine residues in small oligos can be efficiently deaminated in a variety of sequence contexts, some other feature of ssDNA in cells is likely protecting otherwise favorable sites from deamination. For example, DNA-binding proteins could make some deamination sites inaccessible to A3G. In the case of an HIV-1 infection, the HIV-1 nucleocapsid (NC) protein is a known DNA-binding protein that has important functions in the HIV-1 life cycle (63, 82, 130, 150, 159, 217). However, HIV-1 NC and A3G binding is non-competitive on target oligonucleotides, suggesting that NC protein may not prevent access of A3G to a particular target site, and actually could enhance A3G binding (108). Another possibility is that certain HIV-1 ssDNA regions may be protected from A3G due to secondary structure folding that occurs during the reverse transcription process. A3G does not act on dsDNA templates, and therefore ssDNA secondary structure (e.g., stem structures) could act as an accessibility barrier for A3G (227). Also, cytosine bases in small loop structures may also be inaccessible targets, particularly if the proper contacts between enzyme and substrate are no longer in alignment due to physical constraints.

In this study, we have investigated the nature of A3G target sites, in particular the impact of nucleotide bases adjacent to the cytosine dinucleotide target as well as the influence of secondary structure in ssDNAs. Here we demonstrate that by systematic nucleotide base changes on either side of the cytosine dinucleotide target that certain bases are preferred by A3G in order to be optimal targets for cytosine deamination. We

also observed that DNA stems represent poor targets for A3G, and can protect an otherwise desirable target sequence from cytosine deamination. Small loop structures can also protect potential A3G target sequences from cytosine deamination. Taken together, our findings provide the first demonstration that A3G cytosine deamination hotspots are defined by both the sequence context of the cytosine dinucleotide target as well as the ssDNA secondary structure. Such observations provide further information for predicting the locations of cytosine deamination by A3G, which is of particular importance in tracing the origins of HIV-1 genetic variation *in vivo*.

**Results**

*Sequence context as well as ssDNA secondary structure define A3G cytosine deaminase hotspots*

Two sets of ssDNA oligonucleotides were initially tested in which the cytosine dinucleotide was located either in an open (unstructured) location or was located within a structured stem; representative examples of the oligonucleotide structures are shown in Figure 3-1A. The oligonucleotide design strategy helped to minimize sequence changes between the oligonucleotide pairs and distal to the deamination site. However, some additional nucleotide changes were required for some of the oligonucleotides tested in order for the CC dinucleotide to be in the correct structural location within the most stable structure (Table 3-1 and Supplemental Figure 3-1). In particular, the CccC Stem Set 1 has an extra base pair in the stem, though the oligonucleotide sequence is consistent with what is shown in Figure 3-1A. For the TccT Open Set 2 oligonucleotide, the number of loop bases was reduced from 6 to 4 bases, and the number of bases involved in the

stem decreased from 7 to 3 bases. Finally, for the 5 Loop CccC oligonucleotide, the bottom base in the '5 loop' in Figure 2A was changed to a C residue rather than a G residue (Figure 3-2A). The specific predicted mFold structures using the default settings for these three oligonucleotides are indicated in Supplemental Figure 3-1. Other oligonucleotides were tested in which bases on either side of the cytosine dinucleotide were changed. These adjacent bases represent the base locations that are most critical for A3G recognition (182). Cell lysates prepared from A3G-expressing 293 cells or 293 parental cells were used to incubate with each oligonucleotide along with UDG and RNase A as described in the Materials and Methods section. In the presence of A3G cytosine deaminase activity, creation of a uracil base would occur resulting in an abasic site following uracil base excision by UDG. Base hydrolysis of the abasic site would release a FAM signal from the FRET pair.

Figure 3-1B demonstrates that the bases on either side of the cytosine dinucleotide are important for A3G activity when located in a non-structured region. In particular, we observed that adenine, cytosine or thymine bases on either side of the cytosine dinucleotide increased A3G activity ($P < 0.05$) whereas guanine bases on either side of the dinucleotide had a reduced but significant effect ($P < 0.05$). These results indicate that adenine, cytosine or thymine bases on either side of the cytosine dinucleotide enhance A3G activity and guanine bases limit A3G activity.

We further explored the nature of the nucleotide bases on either side of the cytosine dinucleotide by investigating the base preference on 5′ or 3′ side of the dinucleotide site (Figure 3-1C). In particular, when the 5′ base was a guanine, there was little activity detected when the 3′ base was an adenine or thymine, but a high A3G signal

was observed when the 3′ base was a cytosine. When the 5′ base was an adenine, there was moderate A3G activity unless the 3′ base was a cytosine. Finally, when the 5′ base was a thymine, A3G activity was moderate to relatively high when the 3′ base was either a guanine or an adenine, and activity was enhanced if the 3′ base was a cytosine. Significantly reduced A3G activity was observed when cytosine dinucleotides were located within an oligonucleotide stem, indicating that A3G can have difficulty in accessing target bases located in regions in which secondary structure exists, in any sequence context (Figure 3-1C). Taken together, these observations indicate that both the nucleotide base on either side of the cytosine dinucleotide as well as their location in secondary structure can define A3G hotspots.

*Structural constraints in DNA loop bases can limit A3G hotspots*

Given our observation that ssDNA secondary structure can attenuate A3G activity, we further investigated how the location cytosine dinucleotides in ssDNA loop bases could impact A3G activity. To do this, we tested oligonucleotides in which the cytosine dinucleotide was in either a stem bulge or in a ssDNA loop that ranged from 3 bases in size up to 10 bases in size (Figure 3-2A). A3G activity was not affected by the cytosine dinucleotide located in a stem bulge, but had low activity when the dinucleotide was located in a 3 nt base loop where either adenine or cytosine was flanking the cytosine dinucleotide (Figure 3-2B). This indicates that 3 nt base loops can protect A3G hotspots. Interestingly, when cytosine bases flanked the cytosine dinucleotide, high A3G activity was observed within 4–8 nt base loops, but not when adenines flanked the cytosine dinucleotide. Low activity was detected with adenines flanked the cytosine dinucleotide

in the 4 nt base loops, and moderate A3G activity detected when the cytosine dinucleotide was located in 5–7 nt base loops with adenine bases flanking (Figure 3-2B). Higher activity was observed in 8–10 base loops with adenines flanking. This indicates that nucleotide base loop structures can protect cytosine dinucleotides when flanked by adenine bases in seven base or smaller loops. When cytosine bases flank the cytosine dinucleotide, protection is observed only in a 3 nt base loop. Moderate or low A3G activity was observed with thymine or guanine bases flanking the cytosine dinucleotide, respectively. This observation complements the observations made in the absence of secondary structure.

Since the binding of HIV-1 NC and A3G is non-competitive on target oligonucleotides, NC protein may not prevent access of A3G to a particular target site, and may enhance A3G binding (108). It is also formally possible that NC may protect certain HIV-1 ssDNA regions due to secondary structure folding that occurs during reverse transcription. Since A3G does not act on dsDNA templates, ssDNA secondary structure (e.g., stem structures) could act as an accessibility barrier for A3G (227). It is also conceivable that cytosine bases in small loop structures may be inaccessible due to physical constraints. To test for potential affects of HIV-1 NC on A3G activity, we selected an oligonucleotide pair in which there was a clearly significant difference in the FRET signal observed when the CC dinucleotide target was located in either a non-structured or structured region (i.e., AccA set 2 open and stem oligonucleotides; Table 3-1). A HIV-1 NC concentration (i.e., 5 nt per NC) was chosen that is physiologically relevant based upon what is predicted in the virus particle (99, 230). As indicated in Figure 3-3, the addition of NC was found to have no effect on A3G activity when the CC

dinucleotide was either in the AccA set 2 open or stem oligonucleotide.

It has been previously demonstrated that UDG excises uracil residues more efficiently from ssDNA than dsDNA (122), and that the excision of uracil from loops is inefficient. In order to confirm that the FRET signal differences observed with the target cytosine base for cytosine deamination by A3G is in a non-paired region, or in a stem, bulge or DNA loop is actually due to A3G activity and not to UDG, we synthesized oligonucleotides that contain uracil in these different locations. Figure 3-4 shows that UDG is readily able to excise the uracil residue in each of these positions, indicating that the differences that we have observed are due to A3G activity and not due to UDG.

*Experimental confirmation of mFold ssDNA structural predictions*

The oligonucleotides that were used in this study were selected in part based upon their having a single structural prediction in the mFold program (241). Specific parameters have been designed into the mFold program for ssDNA folding that were based upon NMR data generated with ssDNA sequences (7-10, 178, 195). In order to experimentally validate that these predictions for the oligonucleotides used in this study, a small subset were analysed that possessed DNA secondary structures that created restriction enzyme sites. Confirmation of the presence of these restriction sites would provide one line of experimental evidence in support of the structure predicted by the mFold program for oligonucleotide test. Oligo GccG set 1 stem, when folded, creates a Msp I restriction site that is not present in the non-folded version of the oligonucleotide. Figure 3-5A resulted in a strong FRET signal when oligonucleotide GccG set 1 stem was incubated in the presence of Msp I, but not when the non-folding version of the

oligonucleotide (i.e., GccG set 1 open) was incubated with Msp I. This data suggest that the predicted structure for the GccG set 1 stem is correct. We conducted a similar analysis with GccG set 2 stem and GccG set 2 open, where the predicted folded structure for GccG set 2 stem resulted in the creation of an Aci I restriction site, which does not occur in GccG set 2 open (Figure 3-5B). Incubation of each oligonucleotide with Aci I lead to a strong FRET signal only with the GccG set 2 stem oligonucleotide, which also suggests that the structural predictions by mFold for the oligonucleotides used in this study are correct. Although this data support the proposed intramolecular structural predictions, it is formally possible that stable structures could also arise using the set 1 stem or set 2 stem oligonucleotides by the formation of intermolecular homodimers. For instance, the restriction enzyme analysis conducted above with Msp I and Aci I would not be able differentiate per se between a single intramolecular stem versus that of a intermolecular homodimer stem—though stem formation of the participating nucleotide bases would be confirmed.

**Discussion**

The goal of this study was to investigate the determinants for A3G hotspots. To do this, we used an experimental model system in which we used oligonucleotides that were dual-labeled with TAMRA and FAM fluorophores. Lysates from cells stably expressing A3G were incubated with these oligonucleotides, and A3G activity was detected by FRET. Oligonucleotides were designed to test (i) the role of nucleotide bases adjacent to the cytosine dinucleotide target site that is critical for cytosine deamination; and (ii) the role of ssDNA secondary structure, including DNA duplexes, loop sequences

and bulges.

We observed that the ability of A3G to deaminate was found to be greatly dependent upon the nucleotide bases immediately adjacent to the cytosine dinucleotide. Specifically, A3G efficiently deaminates when cytosines, adenines or thymines are adjacent to the cytosine dinucleotide. A3G activity was low when guanine bases were on either side of the cytosine dinucleotide. Previous studies have indicated that A3G prefers a sequence context of 5′-CCCA-3′ or 5′-T/CCC-3′ (21, 234). This corresponds well with the data in our study. In addition, a study has been reported in which 5′-TCCA-3′, 5′-ACCA-3′, and 5′-ACCG-3′ were found to be good substrates for A3G cytosine deaminase activity, whereas 5′-GCCA-3′ and 5′-ACCT-3′ were found to have no or minimal activity, respectively (16). Although these studies support parts of our current study, our observations represent a more extensive and complete study of the preferred bases, and therefore provides greater insight into being able to predict and identify A3G hotspots.

The results from our studies are supported by studies investigating A3G hotspots identified in HIV-1 sequences recovered from infected individuals. Coffin and colleagues found that of the available sites for A3G-mediated cytosine deamination, 40% of 5′-CCCC-3′ sequences, 21% of 5′-ACCA-3′ sequences, 11% of 5′-TCCT-3′ sequences and 0% of 5′-GCCG-3′ sequences were A3G cytosine deamination sites, which is a striking correlation to our data (111). Furthermore, 5′-TCCA-3′, 5′-TCCC-3′ and 5′-ACCT-3′ sequences were found be locations that the authors concluded that A3G-mediated cytosine deamination had occurred, while no cytosine deamination occurred at 5′-GCCT/A-3′ sequences. Finally, HIV-1 proviral sequencing data from our previous

studies of A3G-mediated cytosine deamination of HIV-1 in cell culture found G-to-A mutations (in the positive strand) at either 5′-GGGG-3′, 5′-TGGT-3′ or 5′-TGGA-3′ sequences and few or no mutations at either 5′-AGGA-3′, 5′-AGGT-3′, 5′-TGGC-3′ or 5′-CGGC-3′ sequences (61, 192). Taken together, these data indicate that the A3G hotspot sequence preferences in our cell-free study using oligonucleotides corresponds with strong predictive power to the A3G hotspot preferences observed from HIV-1 proviral DNA sequencing. In addition, A3G mutational hot spot sites are clearly far more complex than the widely cited 5′-CCCA-3′ or 5′-T/CCC-3′ A3G nucleotide sequence preference.

A recent study investigated the features of nucleotide bases that can help define the sequence preference of A3G (182). Exocyclic groups in pyrimidines that are located 1 or 2 nt 5′ of the cytosine targeted by A3G were found to dictate substrate recognition. The exocyclic groups were speculated to be important for stacking or for electrostatic interactions among adjacent bases. When these interactions are disrupted, it was conjectured that it could affect the ability of A3G to recognize the substrate. This hypothesis is supported by our data with the sequence 5′-T/CCCG/A/C-3′. However, we observed that the sequence 5′-ACCA-3′ can also be an efficient target for A3G-mediated cytosine deamination. It is of particular interest to these observations that local sequence context has been found to influence the scanning ability of A3G and that A3G has been proposed to ‘hover’ over 5′-ACCCA-3′ sequences longer than 5′-TCCCT-3′ sequences (201). This observation suggests that additional features of adenine bases may be important in the attraction of A3G to sites of cytosine deamination.

We have also demonstrated in this study that ssDNA secondary structure plays a

vital role in the identification of A3G hotspots. In particular, the data presented here indicate that A3G either has no or low deamination activity for the cytosine dinucleotides located in ssDNA oligonucleotide stem structures or cytosine dinucleotides located in three base loops. The low activity observed on stems may be due to base unzipping from the stem at a low frequency, which would expose the CC dinucleotide. Furthermore, only cytosine dinucleotides flanked by cytosines or adenines were efficiently deaminated when located in ssDNA loops up to 8 bases in size or 8–10 bases respectively, whereas cytosine dinucleotides flanked by adenines resulted in only moderate level of A3G-mediated cytosine deamination in ssDNA oligonucleotide loops 5–7 bases. A related cytosine deaminase family member, AID, has also been reported to be inefficient in deaminating cytosines in ssDNA secondary structures involving stems and loops (125). Since A3G cannot deaminate dsDNA, we hypothesize that ssDNA stems mimic dsDNA in an efficient enough of a manner in order to avoid cytosine deamination. Furthermore, A3G has been previously demonstrated to be processive along ssDNA substrates by sliding and jumping (44, 45). When A3G encounters partially dsDNA, the sliding ability was lost but the jumping ability was retained (44). Therefore, it is tempting to speculate that A3G could 'jump' over stems and loops—and this could help explain at least part of the reduced level of cytosine deamination in those regions. However, the ability of A3G to 'jump' over stems and loops does not account for the differences observed with 5′-CCCC-3′ and 5′-ACCA-3′ sequences in ssDNA loops. While the tortional bend of nucleotides in ssDNA loops may not allow for the proper contacts between A3G and the nucleotide bases, this may be more readily overcome with cytosines rather than adenines. Further studies are needed in order to determine the specific factors behind these

observations. It will be important and beneficial to compare the sequence and secondary structural preferences of other APOBEC3 family members to our findings on APOBEC3G. The addition of HIV-1 NC protein was found in our experiments to have no effect on A3G activity when the CC dinucleotide was either in the AccA set 2 open or stem oligonucleotide. These observations suggest that ssDNA secondary structure (e.g., stem structures) could act as an accessibility barrier for A3G, even in the presence of physiologically relevant concentrations of HIV-1 NC. It is presently unclear how generally applicable these observations are to other CC dinucleotide position locations.

In summary, the observations made in this study provide the first demonstration that A3G cytosine deamination hotspots are defined by both the sequence context of the cytosine dinucleotide target as well as the ssDNA secondary structure. These observations provide useful information for predicting the locations of cytosine deamination by A3G. Such predictions are important for investigations directed at investigating the origins of mutations that are associated with HIV-1 genetic variation (115). Given the high HIV-1 mutation rate (143), and the high rate of G-to-A transition mutations (223), there is intense interest in the origins of mutations that arise during HIV-1 replication. Knowledge on the origins of mutations during HIV-1 replication is important for developing a better understanding of HIV-1 genetic variation and evolution as well as for efforts to purposely elevate HIV-1 mutation to induce lethal mutagenesis (50, 62).

## Materials and Methods

*Preparation of cell lysates*

A cell line stably expressing A3G, 293-A3G clone10 (192) was cultured in DMEM supplemented with 10% FetalClone3 (FC3, Hyclone) serum with 1% penicillin/streptomycin (Invitrogen), and 225 µg/ml neomycin (Invitrogen). The parental 293 cells were maintained in DMEM with 10% FC3, 1% penicillin/streptomycin and 225 µg/ml neomycin. Cells were incubated at 37°C in 5% $CO_2$. Cell lysates for in vitro assays of A3G activity were prepared as previously described (216). Briefly, $5 \times 10^6$ cells (293 or 293-A3G clone 10) were centrifuged at 1000 rpm for 5 min. Cells were resuspended in 250 µl of lysis buffer (0.626% NP40, 10 mM tris-acetate pH 7.4, 50 mM potassium acetate, 100 mM NaCl and 10 mM EDTA) and 50 µl of protease inhibitor cocktail (Cat. # P2714-1BTL, Sigma-Aldrich), and incubated on ice for 15 min. Cell lysates were centrifuged at 1000 rpm for 2 min, and the supernatants transferred to a pre-chilled tube and centrifuged at 16,000 rpm for 10 min. The clarified cell supernatants were then transferred to a new pre-chilled tube and stored at −80°C prior to use.

*Oligonucleotide design and synthesis*

ssDNA oligonucleotide secondary structures were predicted by using mFold with the default DNA settings (http://mfold.rna.albany.edu/?q=mfold/DNA-Folding-Form) (241). The oligonucleotides selected for synthesis had only a single predicted structure and were synthesized dual-labeled with TAMRA and FAM fluorophores (Sigma-Aldrich). Table 3-1 indicates the oligonucleotides used in this study.

*A3G FRET assay*

A fluorescence resonance energy transfer (FRET) based assay was used to detect

cytosine deaminase activity of A3G using DNA oligonucleotides as a substrate using previously described assays with minor modifications (212, 216). Cell lysates were diluted 3:2 in lysis buffer, and 20 µl of the diluted lysates were used per assay using 96 white-welled assay plates (Bio-Rad). A separate solution of 20 pmoles of oligonucleotide, 10 µg RNase A and 0.04 U uracil DNA glycosylase (UDG) were mixed together in 50 mM Tris pH 7.4, 10 mM EDTA buffer and adjusted to a total volume of 50 µl, then transferred to the assay well. The assay plate was then incubated at 37°C for 5 h. Next, 30 µl of 2 M Tris-acetate, pH 7.9 was added to each well and the plate was incubated at 95°C for 2 min and at 4°C for 2.5 min with a CFX96 real-time PCR system (Bio-Rad). The fluorescence was then measured at 4°C. The endpoint fluorescence from the parental 293 cell lysate was subtracted from all experimental samples in order to calculate a relative change in fluorescence due to A3G activity. Experiments were conducted with three independent replicates. For assays involving HIV-1 NC protein (purified NC protein graciously provided by Dr. Rob Gorelick, SAIC, Frederick, MD), experiments were conducted in the presence of HIV-1 NC protein (5 nt per NC) for 1 hr on ice. Following incubation, cell lysates were prepared as described above, added to each well for 5 h at 37°C, and then fluorescence measured.

*Restriction enzyme FRET assay*

To further validate the predicted structures of the FRET oligos, select oligos were digested with restriction enzymes. Regions of the oligo in a stem secondary structure would be cut by the specific restriction enzyme and would release FRET signal. Briefly, 20 pmoles of GccG set 2 open or stem oligo and 5U of Aci I (NEB) were added to a

solution of 1× Buffer 3 (NEB) in 100 ul total volume. For GccG set 1 open or stem oligo, 20 pmoles were added with 2U of MspI (NEB) to a solution of 1× Buffer 4(NEB) in 100 ul total volume. Mixes were added to white-welled 96-well plates (Biorad) and incubated at 37°C for 30 min in a C1000 thermal cycler with a CFX96 real-time system (Biorad). Subsequently, the temperature was adjusted to 4°C for 30 s and the plate was read. Experiments were conducted with three independent replicates.

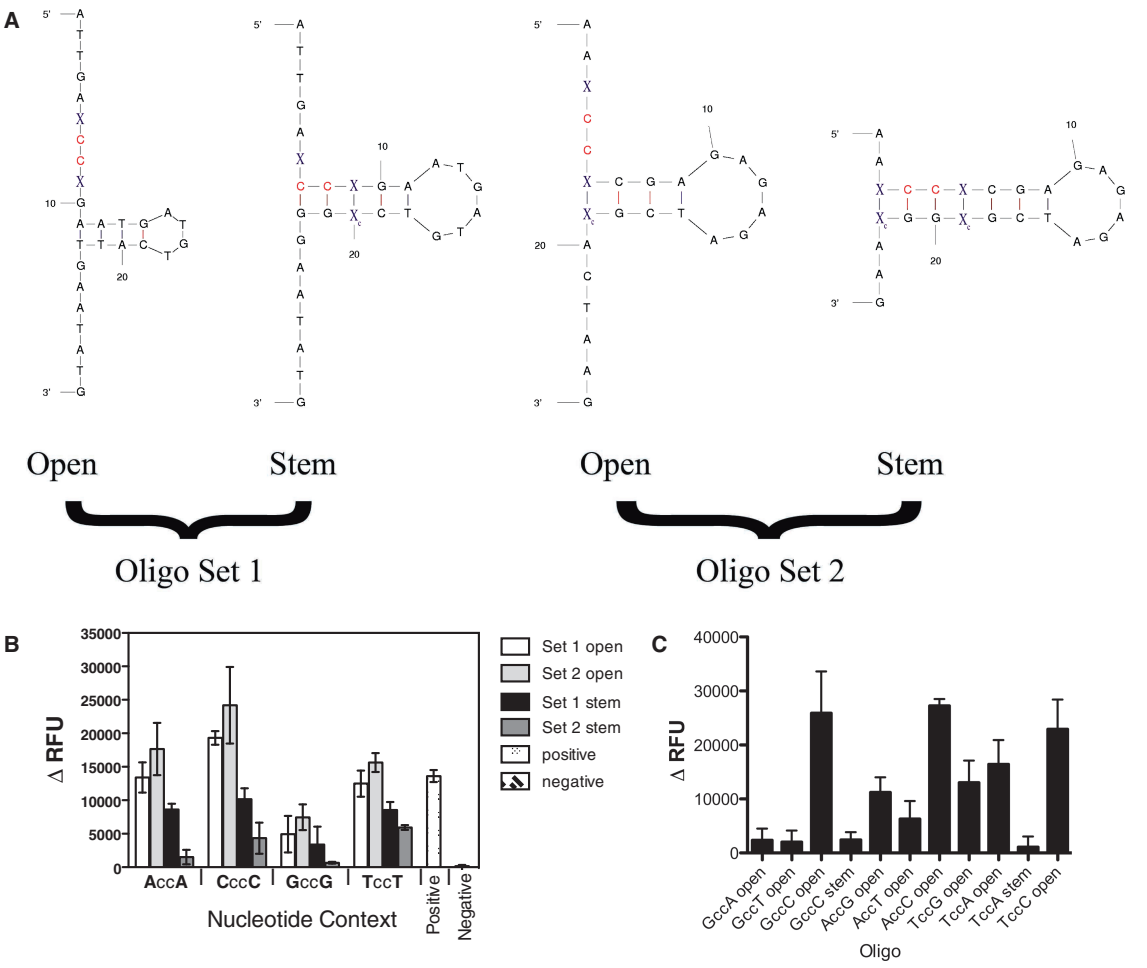



**Figure 3-1. Nucleotide sequence context and ssDNA secondary structure help to define A3G cytosine deaminase hotspots.** (A) Oligonucleotides containing the cytosine dinucleotide targeted by A3G dual-labeled with TAMRA and FAM fluorophores. The red colored 'CC' dinucleotide bases represent the A3G target site. The blue colored 'X' bases represent the positions at which nucleotide bases were changed. The 'open' oligo-nucleotides are defined as the oligonucleotides in which the target cytosine dinucleotide is located in the unstructured region of the ssDNA, and the 'stem' oligonucleotides are defined as the oligonucleotides in which the target cytosine dinucleotide is located within the stem structure. (B) The change in relative fluorescence units (ΔRFU) was calculated

for each experiment by subtracting the RFU from the control 293 cell lysates (baseline negative control) from the 293 cell lysates that stably express A3G. The error bars represent the standard deviation from three independent experiments. The positive control for these experiments was an oligonucleotide previously reported to be cleaved by A3G in an oligonucleotide- based FRET assay (216). (C) The ΔRFU was calculated as described above. The average and standard deviation from three independent experiments is shown.

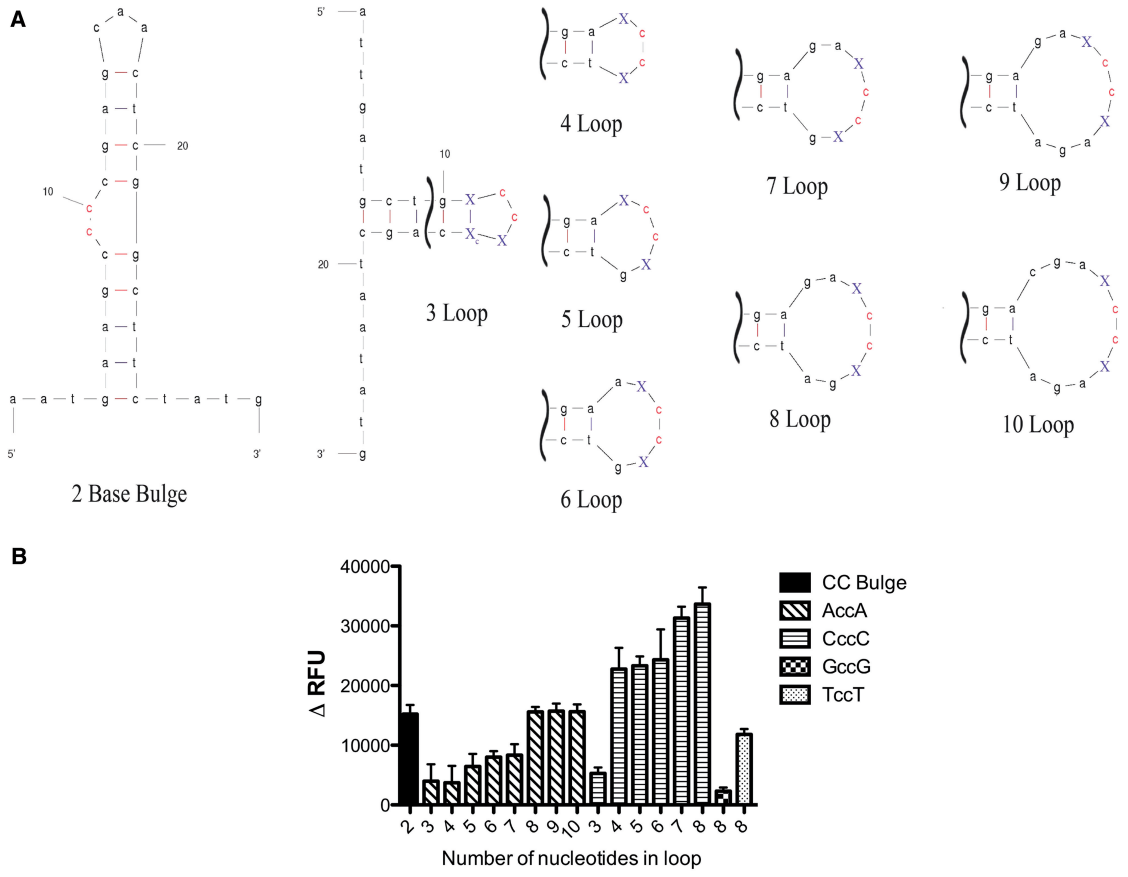



**Figure 3-2. A3G cytosine deaminase activity against a target cytosine dinucloetide is influenced by location in ssDNA base loops but not in a DNA bulge.** (A)

Oligonucleotides used to investigate the influence of ssDNA loop size on A3G activity are shown. The red colored 'CC' dinucleotide bases represent the A3G target site. The blue colored 'X' bases represent the positions at which nucleotide bases were changed. (B) The change in relative fluorescence units (ΔRFU) was calculated for each experiment by subtracting the RFU from the control 293 cell lysates (baseline negative control) from the 293 cell lysates that stably express A3G. The x-axis indicates the number of nucleotide bases in the ssDNA loop. The error bars represent the standard deviation from three independent experiments.

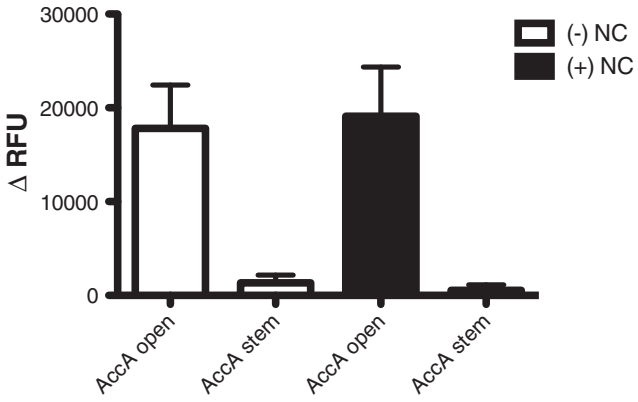



**Figure 3-3. No effect of HIV-1 NC protein on altering the efficiency of A3G deamination.** The AccA set 2 open and stem oligonucleotides were incubated in the presence or absence of HIV-1 NC protein (concentration of 5 nt per NC protein). The change in relative fluorescence units (ΔRFU) was calculated for each experiment by subtracting the RFU from the control 293 cell lysates (baseline negative control) from the 293 cell lysates that stably express A3G. Error bars represent the standard deviation from three independent experiments.

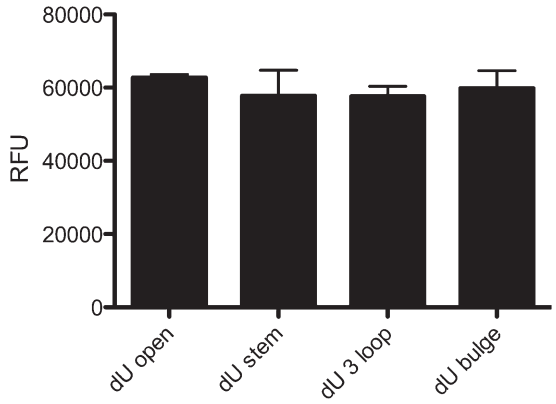



**Figure 3-4. UDG activity is undiminished on ssDNA secondary structures.** The effect of uracil location in oligonucleotides was investigated. Four different oligonucleotides were used in which the target cytosine was replaced with a uracil that was located in a non-paired, stem, bulge or DNA loop region. The relative fluorescent units (RFU) from a uracil in the open, stem, three base loop and bulge location in the presence of UDG is shown. The average and standard deviation from three independent experiments is shown.

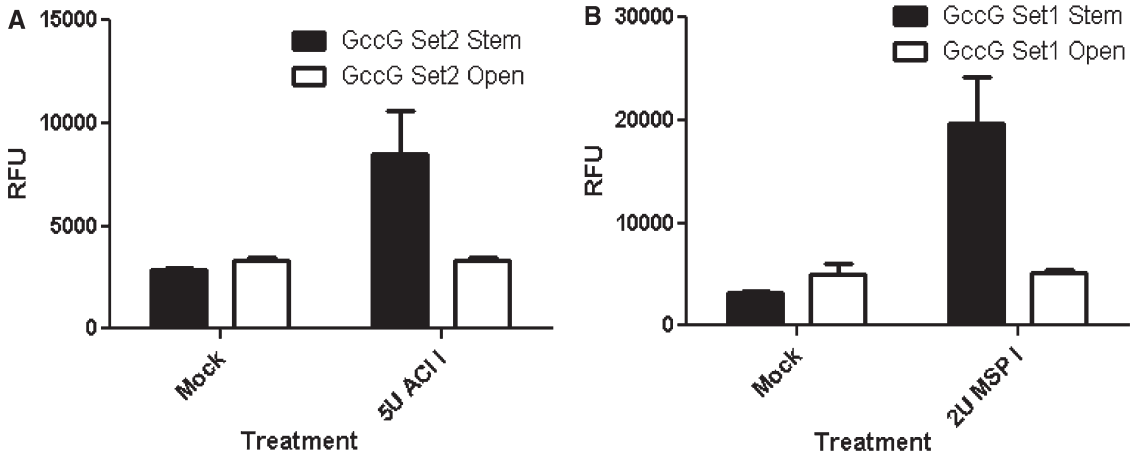
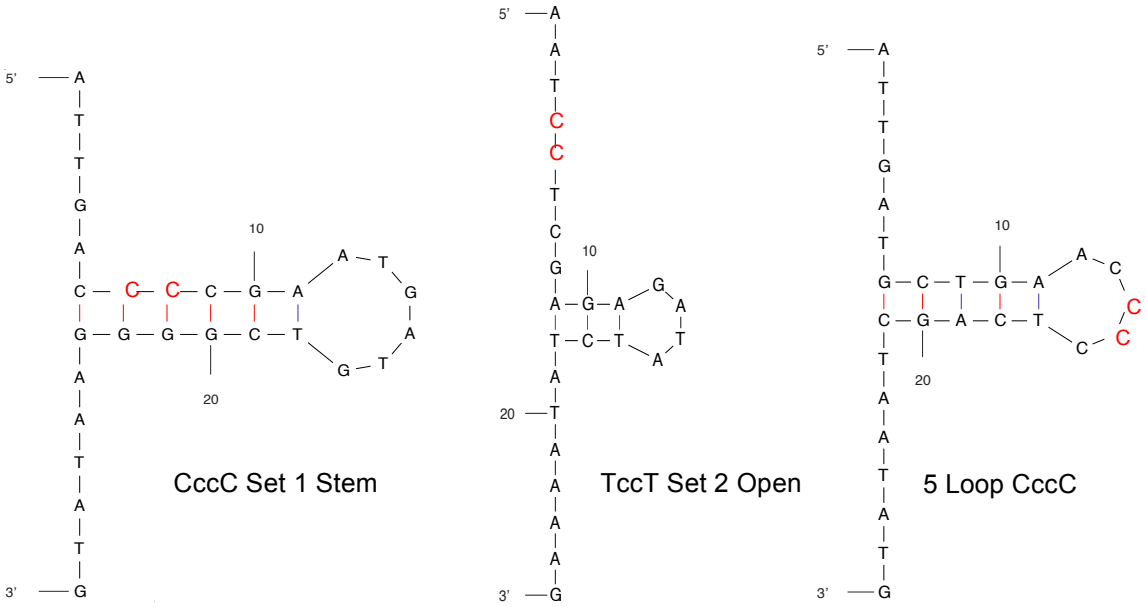



**Figure 3-5. Experimental confirmation of ssDNA secondary structures.** Two sets of

oligonucleotides with restriction enzyme sites in the stem oligonucleotide were used ((A) GCCG Set 2 Stem and GCCG Set 2 Open, and (B) GCCG Set 1 Stem and GCCG Set 1 Open). The stem bases in the structured oligonucleotides create an Aci I restriction site (A) or a Msp I restriction site (B). The oligonucleotides GCCG Set 2 Open and GCCG Set 1 Open did not fold to form the restriction enzyme sites and remained intact. The average and standard deviation from three independent experiments are shown.




**Supplemental Figure 3-1.** Predicted oligonucleotide structures. The most stable predicted structures of oligonucleotides where additional bases changes were needed to maintain the CC dinucleotide in the correct structural location in the most stable structure, as determined by using mFold with the default settings. The structures for the CccC Stem Set 1, TccT Open Set 2, and the 5 Loop CccC oligonucleotides are shown.

**Table 3-1.** Oligonucleotide sequences used in the analysis of the influence of nucleotide sequence and ssDNA secondary structure on the *in vitro* activity of APOBEC3G

| Oligonucleotide | Sequence 5′-3′ | Oligonucleotide | Sequence 5′-3′ |
|---|---|---|---|
| AccA Open Set 1<br>ΔG = −1.90 | ATTGAACCAGAATGATGTCATTGAATATG | AccC Open<br>ΔG = −3.35 | AAACCCCGAGAGAGATCGGACTAAG |
| CccC Open Set 1<br>ΔG = −1.90 | ATTGACCCCGAATGATGTCATTGAATATG | TccA Open<br>ΔG = −2.36 | AATCCACGAGACAGATCGTACTAAG |
| TccT Open Set 1<br>ΔG = −1.90 | ATTGATCCTGAATGATGTCATTGAATATG | TccA Stem<br>ΔG = −6.67 | AATCCACGAGACAGATCGTGGAAAG |
| GccG Open Set 1<br>ΔG = −1.90 | ATTGAGCCGGAATGATGTCATTGAATATG | TccC Open<br>ΔG = −3.35 | AATCCCCGAGAGAGATCGGACTAAG |
| AccA Stem Set 1<br>ΔG = −3.77 | ATTGAACCAGAATGATGTCTGGGAATATG | TccG Open<br>ΔG = −3.47 | AATCCGCGAGAGAGATCGCATCAAG |
| CccC Stem Set 1<br>ΔG = −7.37 | ATTGACCCCGAATGATGTCGGGGAATATG | 3 Loop AccA<br>ΔG = −3.85 | ATTGATGCTGACCATCAGCTAATATG |
| TccT Stem Set 1<br>ΔG = −3.30 | ATTGATCCTGAATGATGTCAGGTAATATG | 3 Loop CccC<br>ΔG = −4.84 | ATTGATGCTGCCCCGCAGCTAATATG |
| GccG Stem Set 1<br>ΔG = −4.81 | ATTGAGCCGGAATGATGTCCGGGAATATG | 4 Loop AccA<br>ΔG = −4.60 | ATTGATGCTGAACCATCAGCTAATATG |
| AccA Open Set 2<br>ΔG = −2.36 | AAACCACGAGAGAGATCGTACTAAG | 4 Loop CccC<br>ΔG = −4.10 | ATTGATGCTGACCCCTCAGCTAATATG |
| CccC Open Set 2<br>ΔG = −3.35 | AACCCCCGAGAGAGATCGGATGAAG | 5 Loop AccA<br>ΔG = −4.60 | ATTGATGCTGAACCAGTCAGCTAATATG |
| TccT Open Set 2<br>ΔG = −2.23 | AATCCTCGAGAGATATCTATAAAAG | 5 Loop CccC<br>ΔG = −4.40 | ATTGATGCTGAACCCCTCAGCTAATATG |
| GccG Open Set 2<br>ΔG = −3.47 | AAGCCGCGAGAGAGATCGCATCAAG | 6 Loop AccA<br>ΔG = −3.90 | ATTGATGCTGAAACCAGTCAGCTAATATG |
| AccA Stem Set 2<br>ΔG = −7.18 | AAACCACGAGAGAGATCGTGGTAAG | 6 Loop CccC<br>ΔG = −3.90 | ATTGATGCTGAACCCCGTCAGCTAATATG |
| CccC Stem Set 2<br>ΔG = −9.31 | AACCCCCGAGAGAGATCGGGGGAAG | 7 Loop CccC<br>ΔG = −3.50 | ATTGATGCTGAGACCCCGTCAGCTAATATG |
| TccT Stem Set 2<br>ΔG = −6.36 | AATCCTCGAGAGAGATCGAGGAAAG | 7 Loop AccA<br>ΔG = −3.50 | ATTGATGCTGAGAACCAGTCAGCTAATATG |
| GccG Stem Set 2<br>ΔG = −9.96 | AAGCCGCGAGAGAGATCGCGGCAAG | 8 Loop AccA<br>ΔG = −3.70 | ATTGATGCTGAGAACCAGATCAGCTAATATG |
| GccA Open<br>ΔG = −4.42 | AAGCCACAAGAGAGATCTTGCTAAG | 8 Loop CccC<br>ΔG = −3.70 | ATTGATGCTGAGACCCCGATCAGCTAATATG |
| GccT Open<br>ΔG = −3.94 | AAGCCTAAAGAGAGATCTTTGAAAG | 8 Loop TccT<br>ΔG = −3.70 | ATTGATGCTGAGATCCTGATCAGCTAATATG |
| GccC Open<br>ΔG = −2.51 | AAGCCCGAAGAGAGATTCGAATAAG | 8 Loop GccG<br>ΔG = −3.70 | ATTGATGCTGAGAGCCGGATCAGCTAATATG |
| GccC Stem<br>ΔG = −8.39 | AAGCCCGAAGAGAGATTCGGGCAAG | 9 Loop AccA<br>ΔG = −3.70 | ATTGATGCTGAGAACCAAGATCAGCTAATATG |
| AccG Open<br>ΔG = −3.47 | AAACCGCGAGAGAGATCGCACTAAG | 10 Loop AccA<br>ΔG = −3.70 | ATTGATGCTGACGAACCAAGATCAGCTAATATG |
| AccT Open<br>ΔG = −1.46 | AAACCTCGAGACAGATCGAACTAAG | dU Open<br>ΔG = −2.67 | AAACUACGAGAGAGATCGTGCTAAG |
| CccC Bulge<br>ΔG = −5.30 | AATGAAGCCCCGAGCAACTCGGCTTCTATG | dU Stem<br>ΔG = −1.72 | ATTGATCCUTGAATGATGTCAGGGGATATG |
| dU Bulge<br>ΔG = −5.30 | AATGAAGCCUCGAGCAACTCGGCTTCTATG | dU 3 Loop<br>ΔG = −3.85 | ATTGATGCTGACUATCAGCTAATATG |

# CHAPTER IV

# DISCUSSION AND FUTURE DIRECTIONS

The studies conducted in this dissertation have led to an increased understanding of the origins of mutations that arise during HIV-1 replication and result in genetic diversity in virus populations. In Chapter II of this dissertation, an investigation of cell type and its influence on HIV-1 diversity was conducted. The results specifically demonstrated that target cell type had a significant influence on HIV-1 vector mutation spectra. In particular, there were significant differences in the frequency of G-to-A and T-to-C transition mutations between the target cell types infected with vector virus. Furthermore, it was observed that numerous G-to-A mutations occurred within the same proviral sequence (i.e., hypermutation). These G-to-A mutations were found to occur at distinct sites (i.e., hotspots) across numerous proviral sequences among the different target cell types analyzed. When these G-to-A hypermutated sequences were removed from the analysis of the virus mutation spectra, there were no longer significant differences in this mutation type among the cell lines. Taken together, these observations suggest that the recovered G-to-A hypermutated proviral sequences are likely attributed to a cellular nucleic acid editing enzyme rather than to HIV-1 reverse transcriptase. However, it should be noted that imbalances in cellular nucleotide pools could also lead to elevated levels of viral mutation, including hypermutation. Nucleic acid editing enzymes, such as the APOBEC3 proteins, mutate specific C residues in the minus-strand DNA during reverse transcription and can lead to C-to-U mutation hotspots (i.e., G-to-A mutation hotspots in the plus-strand DNA). The G-to-A mutations from the hypermutated proviral sequences follow a GA dinucleotide specificity, which provides one line of evidence for involvement of the APOBEC3 proteins. This observation is novel because the G-to-A hypermutation observed in this study was target cell dependent.

This is in contrast to the mechanism of action for APOBEC3 proteins, which have been shown to be dependent on the virus-producing cell. In particular, APOBEC3 proteins have to be packaged into the budding virion within the virus-producing cell in order for subsequent G-to-A hypermutation to be observed.

Future experiments should be directed at investigating the basis for the target cell-associated G-to-A hypermutation phenotype observed. One initial step towards this goal would be to analyze the HIV-1 mutation rate and mutation spectra in the absence of APOBEC3 expression. The analysis from this study has demonstrated that SupT1 cells do not express these proteins. Therefore, HIV-1 vector virus production from SupT1 cells and subsequent infection of fresh SupT1 cells would allow for the analysis of HIV-1 mutation spectra in the absence of APOBEC3 expression. In addition, such an experiment would eliminate expression of APOBEC3 proteins from the virus-producing cell. This would be an important control experiment to begin analysis of a target cell effect responsible for G-to-A hypermutation. Subsequent experiments would then produce HIV-1 vector virus from SupT1 cells and infect CEM-GFP, U373-MAGI, and 293T cells. These experiments would be important control experiments to produce HIV-1 vector virus from cells that do not express APOBEC3 proteins. There are examples of small molecule inhibitors reported in the literature that have the ability to inhibit APOBEC3 protein activity. Conducting experiments in the presence of such an inhibitor would allow for the analysis of the APOBEC3 protein expressing cell lines to determine if differences are observed in the presence of such small molecule inhibitors. Another approach could be the use of siRNAs to knock down APOBEC3 protein expression in both the virus-producing cells as well as the target cells. The experiments described in

Chapter II were conducted with a HIV-1 vector expressing the Vif protein. This should limit if not eliminate the ability of APOBEC3 protein incorporation into virus particles. Experiments with vector virus containing a mutation that prevents Vif protein expression would allow for an analysis of the HIV-1 mutation spectra with viruses that would readily allow for APOBEC3 protein incorporation. This would be a useful control to include with the other experiments described above. Finally, experiments conducted in human primary cells (i.e., primary T-cells and macrophages) as well as comparative studies with human immunodeficiency virus type 2 (HIV-2) would be important extensions of these studies that would further enhance the biological relevance of the observations made in Chapter II.

Present data is limited regarding the role of other cellular enzymes that could result in G-to-A hypermutation. Given this limitation in the literature, experiments investigating other cellular enzymes that could result in the same phenotype are limited. As alluded to above, imbalances in nucleotide pools represents another potential mechanism that could help explain the G-to-A hypermutation phenotype. The central enzyme involved in the creation of nucleotide pools is ribonucleotide reductase. Differential activity in the reduction of NTPs to form dNTPs is an alternative explanation for the differences in G-to-A hypermutation observed among the target cell types investigated in Chapter II. The analysis of dNTP pools and the selective use of ribonucleotide inhibitors could help provide information to determine whether nucleotide pool differences could help explain the data in Chapter II.

The studies described in Chapter III investigated the structural and sequence requirements necessary for the formation of an APOBEC3G hotspot. Since APOBEC3G

can induce sublethal mutations, identifying the sequence and structural features that enable hotspot formation could help identify locations where G-to-A mutations arise that contribute to HIV-1 diversity. The ability to sort out the sources of G-to-A mutations would help to better establish what roles the APOBEC3 proteins play in HIV-1 evolution. Further experiments with sequences from HIV-1 proviruses where G-to-A mutations have been well characterized and are associated with important phenotypes (e.g., HIV-1 drug resistance) would be a logical extension of the studies in Chapter III. Another experimental approach would be to introduce the nucleotide sequence studied in Chapter III into a HIV-1 vector and produce in 293T cells stably overexpressing APOBEC3G in order to analyze in vivo the ability of APOBEC3G to edit in sequences predicted to be in either loop or stem sequences.

The role of the nucleocapsid protein in changing the accessibility of CC dinucleotide sites to APOBEC3G should be further investigated with different oligonucleotides. The studies described in Chapter III with nucleocapsid protein were not exhaustive and it is expected that there would be ssDNA secondary structures that could be altered by nucleocapsid protein and lead to CC dinucleotide sites accessible to the action of APOBEC3G. Such studies should be conducted to further investigate these possibilities.

The studies presented in Chapter III should be extended to other APOBEC3 proteins in order to better understand the influence of ssDNA structure on the ability of these proteins to edit ssDNA templates. APOBEC3G is unique in its CC dinucleotide site specificity and the other family members have a TC dinucleotide specificity. Other oligonucleotides would therefore be needed to analyze the other APOBEC3 proteins. A

key observation made in Chapter III is that APOBEC3G prefers C>A>T>G overall flanking the CC dinucleotide target site. Additionally, 5'- TCCA/G/C-3' results in strong APOBEC3G activity while minimal APOBEC3G activity was observed with 5'-GCCA/T/G-3'. Interestingly, when the target CC site was put in a stem or a 3-base loop, there was little APOBEC3G activity, indicating that ssDNA secondary structure is also important for APOBEC3G recognition. In addition, with A residues flanking the CC in a loop, there was low activity on a 4-base loop and moderate activity on a 5- to 7-base loop. However, with C residues flanking, there was strong activity on a 4- to 8-base loop.

The results from Chapter III reveal the structural and sequence preferences of APOBEC3G that had not been previously reported. These new insights help to better understand the relationship between sequence context, ssDNA structure and APOBEC3G activity. Such information provides key information for predicting sites in the HIV-1 genome where APOBEC3G editing will occur. Such information is useful for determining the influence of APOBEC3G on HIV-1 diversity. Extending such studies to the other APOBEC3 proteins would provide further insight into the sequence context and ssDNA structure that influences site selection for other APOBEC3 proteins. A technical challenge currently is the creation of stable 293T cell lines that overexpress the other APOBEC3 proteins. For the studies in Chapter III, 293T cell lines overexpressing APOBEC3G were already available and characterized. Extending such studies to other APOBEC3 proteins would require the creation of new cell lines. There may be technical challenges due to cell toxicity. If this is the case, the use of transient transfection for creating 293T cells overexpressing a particular APOBEC3 protein of interested should be explored. Ultimately, analysis in cell culture would be important to help to validate

observations made with purified oligonucleotides with experiments conducted in the context of viral replication.

As outlined above, Chapters II and III of this dissertation have described new information regarding the origins of mutations that can influence HIV-1 diversity and evolution. Such basic knowledge provides new insights into the determinants of HIV-1 mutation, which can contribute to HIV-1 diversity and ultimately influence the emergence of HIV-1 drug resistance, immune evasion, viral pathogenesis and virus evolution.

## BIBLIOGRAPHY


1. 2013. *In* US_FDA (ed.).
2. 1981. Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. MMWR Morb Mortal Wkly Rep **30:**305-308.
3. 1981. Pneumocystis pneumonia--Los Angeles. MMWR Morb Mortal Wkly Rep **30:**250-252.
4. 2012. UNAIDS Global Report Fact Sheet.
5. **Albin, J. S., and R. S. Harris.** 2010. Interactions of host APOBEC3 restriction factors with HIV-1 in vivo: implications for therapeutics. Expert Rev Mol Med **12:**e4.
6. **Alce, T. M., and W. Popik.** 2004. APOBEC3G is incorporated into virus-like particles by a direct interaction with HIV-1 Gag nucleocapsid protein. J Biol Chem **279:**34083-34086.
7. **Allawi, H. T., and J. SantaLucia, Jr.** 1998. Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA. Biochemistry **37:**2170-2179.
8. **Allawi, H. T., and J. SantaLucia, Jr.** 1998. Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. Biochemistry **37:**9435-9444.
9. **Allawi, H. T., and J. SantaLucia, Jr.** 1997. Thermodynamics and NMR of internal G.T mismatches in DNA. Biochemistry **36:**10581-10594.
10. **Allawi, H. T., and J. SantaLucia, Jr.** 1998. Thermodynamics of internal C.T mismatches in DNA. Nucleic Acids Res **26:**2694-2701.
11. **Anderson, J. P., R. Daifuku, and L. A. Loeb.** 2004. Viral error catastrophe by mutagenic nucleosides. Annu Rev Microbiol **58:**183-205.
12. **Arribas, M., L. Cabanillas, and E. Lazaro.** 2011. Identification of mutations conferring 5-azacytidine resistance in bacteriophage Qbeta. Virology **417:**343-352.
13. **Baroni, C. D., F. Pezzella, M. Pezzella, B. Macchi, D. Vitolo, S. Uccini, and L. P. Ruco.** 1988. Expression of HIV in lymph node cells of LAS patients. Immunohistology, in situ hybridization, and identification of target cells. Am J Pathol **133:**498-506.
14. **Barre-Sinoussi, F., J. C. Chermann, F. Rey, M. T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vezinet-Brun, C. Rouzioux, W. Rozenbaum, and L. Montagnier.** 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science **220:**868-871.
15. **Bartz, S. R., and M. A. Vodicka.** 1997. Production of high-titer human immunodeficiency virus type 1 pseudotyped with vesicular stomatitis virus glycoprotein. Methods **12:**337-342.
16. **Beale, R. C., S. K. Petersen-Mahrt, I. N. Watt, R. S. Harris, C. Rada, and M. S. Neuberger.** 2004. Comparison of the differential context-dependence of

DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo. J Mol Biol **337:**585-596.

17. **Bebenek, K., J. D. Roberts, and T. A. Kunkel.** 1992. The effects of dNTP pool imbalances on frameshift fidelity during DNA replication. J Biol Chem **267:**3589-3596.
18. **Bender, D. M., J. Bao, A. H. Dantzig, W. D. Diseroad, K. L. Law, N. A. Magnus, J. A. Peterson, E. J. Perkins, Y. J. Pu, S. M. Reutzel-Edens, D. M. Remick, J. J. Starling, G. A. Stephenson, R. K. Vaid, D. Zhang, and J. R. McCarthy.** 2009. Synthesis, crystallization, and biological evaluation of an orally active prodrug of gemcitabine. J Med Chem **52:**6958-6961.
19. **Berkhout, B., and F. J. van Hemert.** 1994. The unusual nucleotide content of the HIV RNA genome results in a biased amino acid composition of HIV proteins. Nucleic Acids Res **22:**1705-1711.
20. **Bishop, K. N., R. K. Holmes, and M. H. Malim.** 2006. Antiviral potency of APOBEC proteins does not correlate with cytidine deamination. J Virol **80:**8450-8458.
21. **Bishop, K. N., R. K. Holmes, A. M. Sheehy, N. O. Davidson, S. J. Cho, and M. H. Malim.** 2004. Cytidine deamination of retroviral DNA by diverse APOBEC proteins. Curr Biol **14:**1392-1396.
22. **Boasso, A., A. W. Hardy, S. A. Anderson, M. J. Dolan, and G. M. Shearer.** 2008. HIV-induced type I interferon and tryptophan catabolism drive T cell dysfunction despite phenotypic activation. PLoS One **3:**e2961.
23. **Boasso, A., J. P. Herbeuval, A. W. Hardy, S. A. Anderson, M. J. Dolan, D. Fuchs, and G. M. Shearer.** 2007. HIV inhibits CD4+ T-cell proliferation by inducing indoleamine 2,3-dioxygenase in plasmacytoid dendritic cells. Blood **109:**3351-3359.
24. **Bogerd, H. P., H. L. Wiegand, B. P. Doehle, K. K. Lueders, and B. R. Cullen.** 2006. APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells. Nucleic Acids Res **34:**89-95.
25. **Bogerd, H. P., H. L. Wiegand, A. E. Hulme, J. L. Garcia-Perez, K. S. O'Shea, J. V. Moran, and B. R. Cullen.** 2006. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. Proc Natl Acad Sci U S A **103:**8780-8785.
26. **Bonnac, L. F., L. M. Mansky, and S. E. Patterson.** 2013. Structure-Activity Relationships and Design of Viral Mutagens and Application to Lethal Mutagenesis. J Med Chem.
27. **Bouchard, J., M. C. Walker, J. M. Leclerc, N. Lapointe, R. Beaulieu, and L. Thibodeau.** 1990. 5-azacytidine and 5-azadeoxycytidine inhibit human immunodeficiency virus type 1 replication in vitro. Antimicrob Agents Chemother **34:**206-209.
28. **Bourara, K., T. J. Liegler, and R. M. Grant.** 2007. Target cell APOBEC3C can induce limited G-to-A mutation in HIV-1. PLoS Pathog **3:**1477-1485.
29. **Boyer, P. L., H. Q. Gao, P. K. Clark, S. G. Sarafianos, E. Arnold, and S. H. Hughes.** 2001. YADD mutants of human immunodeficiency virus type 1 and

Moloney murine leukemia virus reverse transcriptase are resistant to lamivudine triphosphate (3TCTP) in vitro. J Virol **75:**6321-6328.

30. **Brack-Werner, R., A. Kleinschmidt, A. Ludvigsen, W. Mellert, M. Neumann, R. Herrmann, M. C. Khim, A. Burny, N. Muller-Lantzsch, D. Stavrou, and et al.** 1992. Infection of human brain cells by HIV-1: restricted virus production in chronically infected human glial cell lines. Aids **6:**273-285.
31. **Bransteitter, R., P. Pham, M. D. Scharff, and M. F. Goodman.** 2003. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. Proc Natl Acad Sci U S A **100:**4102-4107.
32. **Brenchley, J. M., D. A. Price, T. W. Schacker, T. E. Asher, G. Silvestri, S. Rao, Z. Kazzaz, E. Bornstein, O. Lambotte, D. Altmann, B. R. Blazar, B. Rodriguez, L. Teixeira-Johnson, A. Landay, J. N. Martin, F. M. Hecht, L. J. Picker, M. M. Lederman, S. G. Deeks, and D. C. Douek.** 2006. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. Nat Med **12:**1365-1371.
33. **Brenchley, J. M., T. W. Schacker, L. E. Ruff, D. A. Price, J. H. Taylor, G. J. Beilman, P. L. Nguyen, A. Khoruts, M. Larson, A. T. Haase, and D. C. Douek.** 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. J Exp Med **200:**749-759.
34. **Browne, E. P., C. Allers, and N. R. Landau.** 2009. Restriction of HIV-1 by APOBEC3G is cytidine deaminase-dependent. Virology **387:**313-321.
35. **Bull, J. J., L. A. Meyers, and M. Lachmann.** 2005. Quasispecies made simple. PLoS Comput Biol **1:**e61.
36. **Bull, J. J., R. Sanjuan, and C. O. Wilke.** 2007. Theory of lethal mutagenesis for viruses. J Virol **81:**2930-2939.
37. **Carias, A. M., S. McCoombe, M. McRaven, M. Anderson, N. Galloway, N. Vandergrift, A. J. Fought, J. Lurain, M. Duplantis, R. S. Veazey, and T. J. Hope.** 2013. Defining the interaction of HIV-1 with the mucosal barriers of the female reproductive tract. J Virol **87:**11388-11400.
38. **Carr, B. I., S. Rahbar, Y. Asmeron, A. Riggs, and C. D. Winberg.** 1988. Carcinogenicity and haemoglobin synthesis induction by cytidine analogues. Br J Cancer **57:**395-402.
39. **Casabianca, A., C. Orlandi, A. Fraternale, and M. Magnani.** 2004. Development of a real-time PCR assay using SYBR Green I for provirus load quantification in a murine model of AIDS. J Clin Microbiol **42:**4361-4364.
40. **Cavarelli, M., C. Foglieni, M. Rescigno, and G. Scarlatti.** 2013. R5 HIV-1 envelope attracts dendritic cells to cross the human intestinal epithelium and sample luminal virions via engagement of the CCR5. EMBO Mol Med **5:**776-794.
41. **Cen, S., F. Guo, M. Niu, J. Saadatmand, J. Deflassieux, and L. Kleiman.** 2004. The interaction between HIV-1 Gag and APOBEC3G. J Biol Chem **279:**33177-33184.

42. **Chattopadhyay, S. K., H. C. Morse, 3rd, M. Makino, S. K. Ruscetti, and J. W. Hartley.** 1989. Defective virus is associated with induction of murine retrovirus-induced immunodeficiency syndrome. Proc Natl Acad Sci U S A **86:**3862-3866.
43. **Chattopadhyay, S. K., D. N. Sengupta, T. N. Fredrickson, H. C. Morse, 3rd, and J. W. Hartley.** 1991. Characteristics and contributions of defective, ecotropic, and mink cell focus-inducing viruses involved in a retrovirus-induced immunodeficiency syndrome of mice. J Virol **65:**4232-4241.
44. **Chelico, L., P. Pham, P. Calabrese, and M. F. Goodman.** 2006. APOBEC3G DNA deaminase acts processively 3' --> 5' on single-stranded DNA. Nat Struct Mol Biol **13:**392-399.
45. **Chelico, L., E. J. Sacho, D. A. Erie, and M. F. Goodman.** 2008. A model for oligomeric regulation of APOBEC3G cytosine deaminase-dependent restriction of HIV. J Biol Chem **283:**13780-13791.
46. **Chen, H., C. E. Lilley, Q. Yu, D. V. Lee, J. Chou, I. Narvaiza, N. R. Landau, and M. D. Weitzman.** 2006. APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. Curr Biol **16:**480-485.
47. **Chiu, Y. L., and W. C. Greene.** 2008. The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. Annu Rev Immunol **26:**317-353.
48. **Chiu, Y. L., H. E. Witkowska, S. C. Hall, M. Santiago, V. B. Soros, C. Esnault, T. Heidmann, and W. C. Greene.** 2006. High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. Proc Natl Acad Sci U S A **103:**15588-15593.
49. **Clouser, C. L., C. M. Holtz, M. Mullett, D. L. Crankshaw, J. E. Briggs, J. Chauhan, I. M. VanHoutan, S. E. Patterson, and L. M. Mansky.** 2011. Analysis of the ex vivo and in vivo antiretroviral activity of gemcitabine. PLoS One **6:**e15840.
50. **Clouser, C. L., S. E. Patterson, and L. M. Mansky.** 2010. Exploiting drug repositioning for discovery of a novel HIV combination therapy. J Virol **84:**9301-9309.
51. **Codoner, F. M., J. A. Daros, R. V. Sole, and S. F. Elena.** 2006. The fittest versus the flattest: experimental confirmation of the quasispecies effect with subviral pathogens. PLoS Pathog **2:**e136.
52. **Coffin, J. M.** 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science **267:**483-489.
53. **Contreras, A. M., Y. Hiasa, W. He, A. Terella, E. V. Schmidt, and R. T. Chung.** 2002. Viral RNA mutations are region specific and increased by ribavirin in a full-length hepatitis C virus replication system. J Virol **76:**8505-8517.
54. **Crotty, S., C. E. Cameron, and R. Andino.** 2001. RNA virus error catastrophe: direct molecular test by using ribavirin. Proc Natl Acad Sci U S A **98:**6895-6900.

55. **Crotty, S., D. Maag, J. J. Arnold, W. Zhong, J. Y. Lau, Z. Hong, R. Andino, and C. E. Cameron.** 2000. The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. Nat Med **6:**1375-1379.
56. **Cuevas, J. M., F. Gonzalez-Candelas, A. Moya, and R. Sanjuan.** 2009. Effect of ribavirin on the mutation rate and spectrum of hepatitis C virus in vivo. J Virol **83:**5760-5764.
57. **Cupples, C. G., and J. H. Miller.** 1989. A set of lacZ mutations in Escherichia coli that allow rapid detection of each of the six base substitutions. Proc Natl Acad Sci U S A **86:**5345-5349.
58. **Dang, Y., X. Wang, W. J. Esselman, and Y. H. Zheng.** 2006. Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. J Virol **80:**10522-10533.
59. **Dapp, M. J., C. L. Clouser, S. Patterson, and L. M. Mansky.** 2009. 5-Azacytidine can induce lethal mutagenesis in human immunodeficiency virus type 1. J Virol **83:**11950-11958.
60. **Dapp, M. J., R. H. Heineman, and L. M. Mansky.** 2013. Interrelationship between HIV-1 fitness and mutation rate. J Mol Biol **425:**41-53.
61. **Dapp, M. J., C. M. Holtz, and L. M. Mansky.** 2012. Concomitant lethal mutagenesis of human immunodeficiency virus type 1. J Mol Biol **419:**158-170.
62. **Dapp, M. J., S. E. Patterson, and L. M. Mansky.** 2013. Back to the future: revisiting HIV-1 lethal mutagenesis. Trends Microbiol **21:**56-62.
63. **Darlix, J. L., J. Godet, R. Ivanyi-Nagy, P. Fosse, O. Mauffret, and Y. Mely.** 2011. Flexible nature and specific functions of the HIV-1 nucleocapsid protein. J Mol Biol **410:**565-581.
64. **Delebecque, F., R. Suspene, S. Calattini, N. Casartelli, A. Saib, A. Froment, S. Wain-Hobson, A. Gessain, J. P. Vartanian, and O. Schwartz.** 2006. Restriction of foamy viruses by APOBEC cytidine deaminases. J Virol **80:**605-614.
65. **Di Noia, J. M., and M. S. Neuberger.** 2007. Molecular mechanisms of antibody somatic hypermutation. Annu Rev Biochem **76:**1-22.
66. **Diamond, T. L., M. Roshal, V. K. Jamburuthugoda, H. M. Reynolds, A. R. Merriam, K. Y. Lee, M. Balakrishnan, R. A. Bambara, V. Planelles, S. Dewhurst, and B. Kim.** 2004. Macrophage tropism of HIV-1 depends on efficient cellular dNTP utilization by reverse transcriptase. J Biol Chem **279:**51545-51553.
67. **Dias, A. S., M. J. Bester, R. F. Britz, and Z. Apostolides.** 2006. Animal models used for the evaluation of antiretroviral therapies. Curr HIV Res **4:**431-446.
68. **Doehle, B. P., A. Schafer, H. L. Wiegand, H. P. Bogerd, and B. R. Cullen.** 2005. Differential sensitivity of murine leukemia virus to APOBEC3-mediated inhibition is governed by virion exclusion. J Virol **79:**8201-8207.
69. **Domingo, E.** 2000. Viruses at the edge of adaptation. Virology **270:**251-253.
70. **Domingo, E., C. Escarmis, E. Lazaro, and S. C. Manrubia.** 2005. Quasispecies dynamics and RNA virus extinction. Virus Res **107:**129-139.

71. **Domingo, E., and J. J. Holland.** 1997. RNA virus mutations and fitness for survival. Annu Rev Microbiol **51:**151-178.
72. **Doria, M., F. Neri, A. Gallo, M. G. Farace, and A. Michienzi.** 2009. Editing of HIV-1 RNA by the double-stranded RNA deaminase ADAR1 stimulates viral infection. Nucleic Acids Res **37:**5848-5858.
73. **Doria, M., S. Tomaselli, F. Neri, S. A. Ciafre, M. G. Farace, A. Michienzi, and A. Gallo.** 2011. ADAR2 editing enzyme is a novel human immunodeficiency virus-1 proviral factor. J Gen Virol **92:**1228-1232.
74. **Dorweiler, I. J., S. J. Ruone, H. Wang, R. W. Burry, and L. M. Mansky.** 2006. Role of the human T-cell leukemia virus type 1 PTAP motif in Gag targeting and particle release. J Virol **80:**3634-3643.
75. **Dutko, J. A., A. Schafer, A. E. Kenny, B. R. Cullen, and M. J. Curcio.** 2005. Inhibition of a yeast LTR retrotransposon by human APOBEC3 cytidine deaminases. Curr Biol **15:**661-666.
76. **Eckstein, D. A., M. L. Penn, Y. D. Korin, D. D. Scripture-Adams, J. A. Zack, J. F. Kreisberg, M. Roederer, M. P. Sherman, P. S. Chin, and M. A. Goldsmith.** 2001. HIV-1 actively replicates in naive CD4(+) T cells residing within human lymphoid tissues. Immunity **15:**671-682.
77. **Eigen, M.** 1993. Viral quasispecies. Sci Am **269:**42-49.
78. **Elion, G. B., S. Singer, and G. H. Hitchings.** 1954. Antagonists of nucleic acid derivatives. VIII. Synergism in combinations of biochemically related antimetabolites. J Biol Chem **208:**477-488.
79. **Gallo, R. C., P. S. Sarin, E. P. Gelmann, M. Robert-Guroff, E. Richardson, V. S. Kalyanaraman, D. Mann, G. D. Sidhu, R. E. Stahl, S. Zolla-Pazner, J. Leibowitch, and M. Popovic.** 1983. Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). Science **220:**865-867.
80. **Giorgi, J. V., L. E. Hultin, J. A. McKeating, T. D. Johnson, B. Owens, L. P. Jacobson, R. Shih, J. Lewis, D. J. Wiley, J. P. Phair, S. M. Wolinsky, and R. Detels.** 1999. Shorter survival in advanced human immunodeficiency virus type 1 infection is more closely associated with T lymphocyte activation than with plasma virus burden or virus chemokine coreceptor usage. J Infect Dis **179:**859-870.
81. **Glick, M., M. Trope, O. Bagasra, and M. E. Pliskin.** 1991. Human immunodeficiency virus infection of fibroblasts of dental pulp in seropositive patients. Oral Surg Oral Med Oral Pathol **71:**733-736.
82. **Godet, J., and Y. Mely.** 2010. Biophysical studies of the nucleic acid chaperone properties of the HIV-1 nucleocapsid protein. RNA Biol **7:**687-699.
83. **Goila-Gaur, R., and K. Strebel.** 2008. HIV-1 Vif, APOBEC, and intrinsic immunity. Retrovirology **5:**51.
84. **Graci, J. D., and C. E. Cameron.** 2008. Therapeutically targeting RNA viruses via lethal mutagenesis. Future Virol **3:**553-566.
85. **Graci, J. D., D. A. Harki, V. S. Korneeva, J. P. Edathil, K. Too, D. Franco, E. D. Smidansky, A. V. Paul, B. R. Peterson, D. M. Brown, D. Loakes, and C. E.**

**Cameron.** 2007. Lethal mutagenesis of poliovirus mediated by a mutagenic pyrimidine analogue. J Virol **81:**11256-11266.
86. **Guadalupe, M., E. Reay, S. Sankaran, T. Prindiville, J. Flamm, A. McNeil, and S. Dandekar.** 2003. Severe CD4+ T-cell depletion in gut lymphoid tissue during primary human immunodeficiency virus type 1 infection and substantial delay in restoration following highly active antiretroviral therapy. J Virol **77:**11708-11717.
87. **Guo, F., S. Cen, M. Niu, J. Saadatmand, and L. Kleiman.** 2006. Inhibition of formula-primed reverse transcription by human APOBEC3G during human immunodeficiency virus type 1 replication. J Virol **80:**11710-11722.
88. **Gurney, K. B., J. Elliott, H. Nassanian, C. Song, E. Soilleux, I. McGowan, P. A. Anton, and B. Lee.** 2005. Binding and transfer of human immunodeficiency virus by DC-SIGN+ cells in human rectal mucosa. J Virol **79:**5762-5773.
89. **Hache, G., M. T. Liddament, and R. S. Harris.** 2005. The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. J Biol Chem **280:**10920-10924.
90. **Hache, G., L. M. Mansky, and R. S. Harris.** 2006. Human APOBEC3 proteins, retrovirus restriction, and HIV drug resistance. AIDS Rev **8:**148-157.
91. **Hall, M. J., R. F. Middleton, and D. Westmacott.** 1983. The fractional inhibitory concentration (FIC) index as a measure of synergy. J Antimicrob Chemother **11:**427-433.
92. **Hankins, C. A., and M. R. Dybul.** 2013. The promise of pre-exposure prophylaxis with antiretroviral drugs to prevent HIV transmission: a review. Curr Opin HIV AIDS **8:**50-58.
93. **Harrington, R. D., and A. P. Geballe.** 1993. Cofactor requirement for human immunodeficiency virus type 1 entry into a CD4-expressing human cell line. J Virol **67:**5939-5947.
94. **Harris, K. S., W. Brabant, S. Styrchak, A. Gall, and R. Daifuku.** 2005. KP-1212/1461, a nucleoside designed for the treatment of HIV by viral mutagenesis. Antiviral Res **67:**1-9.
95. **Harris, R. S., K. N. Bishop, A. M. Sheehy, H. M. Craig, S. K. Petersen-Mahrt, I. N. Watt, M. S. Neuberger, and M. H. Malim.** 2003. DNA deamination mediates innate immunity to retroviral infection. Cell **113:**803-809.
96. **Harris, R. S., and M. T. Liddament.** 2004. Retroviral restriction by APOBEC proteins. Nat Rev Immunol **4:**868-877.
97. **Hartley, J. W., T. N. Fredrickson, R. A. Yetter, M. Makino, and H. C. Morse, 3rd.** 1989. Retrovirus-induced murine acquired immunodeficiency syndrome: natural history of infection and differing susceptibility of inbred mouse strains. J Virol **63:**1223-1231.
98. **Heinemann, V., Y. Z. Xu, S. Chubb, A. Sen, L. W. Hertel, G. B. Grindey, and W. Plunkett.** 1990. Inhibition of ribonucleotide reduction in CCRF-CEM cells by 2',2'-difluorodeoxycytidine. Mol Pharmacol **38:**567-572.
99. **Henderson, L. E., M. A. Bowers, R. C. Sowder, 2nd, S. A. Serabyn, D. G. Johnson, J. W. Bess, Jr., L. O. Arthur, D. K. Bryant, and C. Fenselau.** 1992.

Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: posttranslational modifications, proteolytic processings, and complete amino acid sequences. J Virol **66:**1856-1865.

100. **Herbeuval, J. P., J. C. Grivel, A. Boasso, A. W. Hardy, C. Chougnet, M. J. Dolan, H. Yagita, J. D. Lifson, and G. M. Shearer.** 2005. CD4+ T-cell death induced by infectious and noninfectious HIV-1: role of type 1 interferon-dependent, TRAIL/DR5-mediated apoptosis. Blood **106:**3524-3531.
101. **Herbeuval, J. P., A. W. Hardy, A. Boasso, S. A. Anderson, M. J. Dolan, M. Dy, and G. M. Shearer.** 2005. Regulation of TNF-related apoptosis-inducing ligand on primary CD4+ T cells by HIV-1: role of type I IFN-producing plasmacytoid dendritic cells. Proc Natl Acad Sci U S A **102:**13974-13979.
102. **Herbeuval, J. P., J. Nilsson, A. Boasso, A. W. Hardy, M. J. Kruhlak, S. A. Anderson, M. J. Dolan, M. Dy, J. Andersson, and G. M. Shearer.** 2006. Differential expression of IFN-alpha and TRAIL/DR5 in lymphoid tissue of progressor versus nonprogressor HIV-1-infected patients. Proc Natl Acad Sci U S A **103:**7000-7005.
103. **Holland, J. J., E. Domingo, J. C. de la Torre, and D. A. Steinhauer.** 1990. Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. J Virol **64:**3960-3962.
104. **Holmes, E. C.** 2003. Error thresholds and the constraints to RNA virus evolution. Trends Microbiol **11:**543-546.
105. **Hsiou, Y., J. Ding, K. Das, A. D. Clark, Jr., S. H. Hughes, and E. Arnold.** 1996. Structure of unliganded HIV-1 reverse transcriptase at 2.7 A resolution: implications of conformational changes for polymerization and inhibition mechanisms. Structure **4:**853-860.
106. **Hultquist, J. F., J. A. Lengyel, E. W. Refsland, R. S. LaRue, L. Lackey, W. L. Brown, and R. S. Harris.** 2011. Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. J Virol **85:**11220-11234.
107. **Issa, J. P., and H. Kantarjian.** 2005. Azacitidine. Nat Rev Drug Discov **Suppl:**S6-7.
108. **Iwatani, Y., H. Takeuchi, K. Strebel, and J. G. Levin.** 2006. Biochemical activities of highly purified, catalytically active human APOBEC3G: correlation with antiviral effect. J Virol **80:**5992-6002.
109. **Jackson-Grusby, L., P. W. Laird, S. N. Magge, B. J. Moeller, and R. Jaenisch.** 1997. Mutagenicity of 5-aza-2'-deoxycytidine is mediated by the mammalian DNA methyltransferase. Proc Natl Acad Sci U S A **94:**4681-4685.
110. **Jarmuz, A., A. Chester, J. Bayliss, J. Gisbourne, I. Dunham, J. Scott, and N. Navaratnam.** 2002. An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. Genomics **79:**285-296.
111. **Jern, P., R. A. Russell, V. K. Pathak, and J. M. Coffin.** 2009. Likely role of APOBEC3G-mediated G-to-A mutations in HIV-1 evolution and drug resistance. PLoS Pathog **5:**e1000367.

112. **Jolicoeur, P.** 1991. Murine acquired immunodeficiency syndrome (MAIDS): an animal model to study the AIDS pathogenesis. Faseb J **5:**2398-2405.
113. **Julias, J. G., and V. K. Pathak.** 1998. Deoxyribonucleoside triphosphate pool imbalances in vivo are associated with an increased retroviral mutation rate. J Virol **72:**7941-7949.
114. **Kaminskas, E., A. T. Farrell, Y. C. Wang, R. Sridhara, and R. Pazdur.** 2005. FDA drug approval summary: azacitidine (5-azacytidine, Vidaza) for injectable suspension. Oncologist **10:**176-182.
115. **Kijak, G. H., M. Janini, S. Tovanabutra, E. E. Sanders-Buell, D. L. Birx, M. L. Robb, N. L. Michael, and F. E. McCutchan.** 2007. HyperPack: a software package for the study of levels, contexts, and patterns of APOBEC-mediated hypermutation in HIV. AIDS Res Hum Retroviruses **23:**554-557.
116. **Kim, E. Y., T. Bhattacharya, K. Kunstman, P. Swantek, F. A. Koning, M. H. Malim, and S. M. Wolinsky.** 2010. Human APOBEC3G-mediated editing can promote HIV-1 sequence diversification and accelerate adaptation to selective pressure. J Virol **84:**10402-10405.
117. **Kinomoto, M., T. Kanno, M. Shimura, Y. Ishizaka, A. Kojima, T. Kurata, T. Sata, and K. Tokunaga.** 2007. All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. Nucleic Acids Res **35:**2955-2964.
118. **Kobayashi, M., A. Takaori-Kondo, K. Shindo, A. Abudu, K. Fukunaga, and T. Uchiyama.** 2004. APOBEC3G targets specific virus species. J Virol **78:**8238-8244.
119. **Kohli, R. M., R. W. Maul, A. F. Guminski, R. L. McClure, K. S. Gajula, H. Saribasak, M. A. McMahon, R. F. Siliciano, P. J. Gearhart, and J. T. Stivers.** 2010. Local sequence targeting in the AID/APOBEC family differentially impacts retroviral restriction and antibody diversification. J Biol Chem **285:**40956-40964.
120. **Kohlstaedt, L. A., J. Wang, J. M. Friedman, P. A. Rice, and T. A. Steitz.** 1992. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. Science **256:**1783-1790.
121. **Koning, F. A., C. Goujon, H. Bauby, and M. H. Malim.** 2011. Target cell-mediated editing of HIV-1 cDNA by APOBEC3 proteins in human macrophages. J Virol **85:**13448-13452.
122. **Kumar, N. V., and U. Varshney.** 1994. Inefficient excision of uracil from loop regions of DNA oligomers by E. coli uracil DNA glycosylase. Nucleic Acids Res **22:**3737-3741.
123. **Lanford, R. E., D. Chavez, B. Guerra, J. Y. Lau, Z. Hong, K. M. Brasky, and B. Beames.** 2001. Ribavirin induces error-prone replication of GB virus B in primary tamarin hepatocytes. J Virol **75:**8074-8081.
124. **Langlois, M. A., R. C. Beale, S. G. Conticello, and M. S. Neuberger.** 2005. Mutational comparison of the single-domained APOBEC3C and double-domained APOBEC3F/G anti-retroviral cytidine deaminases provides insight into their DNA target site specificities. Nucleic Acids Res **33:**1913-1923.

125. **Larijani, M., and A. Martin.** 2007. Single-stranded DNA structure and positional context of the target cytidine determine the enzymatic efficiency of AID. Mol Cell Biol **27:**8038-8048.
126. **Lavelle, D., J. Chin, K. Vaitkus, S. Redkar, P. Phiasivongsa, C. Tang, R. Will, M. Hankewych, B. Roxas, M. Singh, Y. Saunthararajah, and J. Desimone.** 2007. Oral decitabine reactivates expression of the methylated gamma-globin gene in Papio anubis. Am J Hematol **82:**981-985.
127. **Lecossier, D., F. Bouchonnet, F. Clavel, and A. J. Hance.** 2003. Hypermutation of HIV-1 DNA in the absence of the Vif protein. Science **300:**1112.
128. **Lee, L. M., J. M. Karon, R. Selik, J. J. Neal, and P. L. Fleming.** 2001. Survival after AIDS diagnosis in adolescents and adults during the treatment era, United States, 1984-1997. Jama **285:**1308-1315.
129. **Levi, L. I., N. F. Gnadig, S. Beaucourt, M. J. McPherson, B. Baron, J. J. Arnold, and M. Vignuzzi.** 2010. Fidelity variants of RNA dependent RNA polymerases uncover an indirect, mutagenic activity of amiloride compounds. PLoS Pathog **6:**e1001163.
130. **Levin, J. G., M. Mitra, A. Mascarenhas, and K. Musier-Forsyth.** 2010. Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription. RNA Biol **7:**754-774.
131. **Li, L. H., E. J. Olin, H. H. Buskirk, and L. M. Reineke.** 1970. Cytotoxicity and mode of action of 5-azacytidine on L1210 leukemia. Cancer Res **30:**2760-2769.
132. **Loeb, L. A., J. M. Essigmann, F. Kazazi, J. Zhang, K. D. Rose, and J. I. Mullins.** 1999. Lethal mutagenesis of HIV with mutagenic nucleoside analogs. Proc Natl Acad Sci U S A **96:**1492-1497.
133. **Lombardi, V. C., F. W. Ruscetti, J. Das Gupta, M. A. Pfost, K. S. Hagen, D. L. Peterson, S. K. Ruscetti, R. K. Bagni, C. Petrow-Sadowski, B. Gold, M. Dean, R. H. Silverman, and J. A. Mikovits.** 2009. Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome. Science **326:**585-589.
134. **Longerich, S., U. Basu, F. Alt, and U. Storb.** 2006. AID in somatic hypermutation and class switch recombination. Curr Opin Immunol **18:**164-174.
135. **Lori, F., and J. Lisziewicz.** 1999. Hydroxyurea: overview of clinical data and antiretroviral and immunomodulatory effects. Antivir Ther **4 Suppl 3:**101-108.
136. **Lutchman, G., S. Danehower, B. C. Song, T. J. Liang, J. H. Hoofnagle, M. Thomson, and M. G. Ghany.** 2007. Mutation rate of the hepatitis C virus NS5B in patients undergoing treatment with ribavirin monotherapy. Gastroenterology **132:**1757-1766.
137. **Maksakova, I. A., M. T. Romanish, L. Gagnier, C. A. Dunn, L. N. van de Lagemaat, and D. L. Mager.** 2006. Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. PLoS Genet **2:**e2.

138. **Malim, M. H.** 2009. APOBEC proteins and intrinsic resistance to HIV-1 infection. Philos Trans R Soc Lond B Biol Sci **364:**675-687.
139. **Malim, M. H., and M. Emerman.** 2008. HIV-1 accessory proteins--ensuring viral survival in a hostile environment. Cell Host Microbe **3:**388-398.
140. **Mangeat, B., P. Turelli, G. Caron, M. Friedli, L. Perrin, and D. Trono.** 2003. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature **424:**99-103.
141. **Mansky, L. M.** 1996. Forward mutation rate of human immunodeficiency virus type 1 in a T lymphoid cell line. AIDS Res Hum Retroviruses **12:**307-314.
142. **Mansky, L. M.** 1998. Retrovirus mutation rates and their role in genetic variation. J Gen Virol **79 ( Pt 6):**1337-1345.
143. **Mansky, L. M., and H. M. Temin.** 1995. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. J Virol **69:**5087-5094.
144. **Mariani, R., D. Chen, B. Schrofelbauer, F. Navarro, R. Konig, B. Bollman, C. Munk, H. Nymark-McMahon, and N. R. Landau.** 2003. Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif. Cell **114:**21-31.
145. **Mayhew, C., O. Oakley, J. Piper, N. K. Hughes, J. Phillips, N. J. Birch, H. L. Elford, and V. S. Gallicchio.** 1997. Effective use of ribonucleotide reductase inhibitors (Didox and Trimidox) alone or in combination with didanosine (ddI) to suppress disease progression and increase survival in murine acquired immunodeficiency syndrome (MAIDS). Cell Mol Biol (Noisy-le-grand) **43:**1019-1029.
146. **Mayhew, C. N., L. J. Mampuru, D. Chendil, M. M. Ahmed, J. D. Phillips, R. N. Greenberg, H. L. Elford, and V. S. Gallicchio.** 2002. Suppression of retrovirus-induced immunodeficiency disease (murine AIDS) by trimidox and didox: novel ribonucleotide reductase inhibitors with less bone marrow toxicity than hydroxyurea. Antiviral Res **56:**167-181.
147. **Mbisa, J. L., R. Barr, J. A. Thomas, N. Vandegraaff, I. J. Dorweiler, E. S. Svarovskaia, W. L. Brown, L. M. Mansky, R. J. Gorelick, R. S. Harris, A. Engelman, and V. K. Pathak.** 2007. Human immunodeficiency virus type 1 cDNAs produced in the presence of APOBEC3G exhibit defects in plus-strand DNA transfer and integration. J Virol **81:**7099-7110.
148. **Mehandru, S., M. A. Poles, K. Tenner-Racz, A. Horowitz, A. Hurley, C. Hogan, D. Boden, P. Racz, and M. Markowitz.** 2004. Primary HIV-1 infection is associated with preferential depletion of CD4+ T lymphocytes from effector sites in the gastrointestinal tract. J Exp Med **200:**761-770.
149. **Mezei, M., and J. Minarovits.** 2006. Reversal of HIV drug resistance and novel strategies to curb HIV infection: the viral infectivity factor Vif as a target and tool of therapy. Curr Drug Targets **7:**881-885.
150. **Mirambeau, G., S. Lyonnais, and R. J. Gorelick.** 2010. Features, processing states, and heterologous protein interactions in the modulation of the retroviral nucleocapsid protein function. RNA Biol **7:**724-734.

151. **Momparler, R. L.** 2005. Pharmacology of 5-Aza-2'-deoxycytidine (decitabine). Semin Hematol **42:**S9-16.
152. **Momparler, R. L., D. Y. Bouffard, L. F. Momparler, J. Dionne, K. Belanger, and J. Ayoub.** 1997. Pilot phase I-II study on 5-aza-2'-deoxycytidine (Decitabine) in patients with metastatic lung cancer. Anticancer Drugs **8:**358-368.
153. **Momparler, R. L., J. Samson, L. F. Momparler, and G. E. Rivard.** 1984. Cell cycle effects and cellular pharmacology of 5-aza-2'-deoxycytidine. Cancer Chemother Pharmacol **13:**191-194.
154. **Morse, H. C., 3rd, S. K. Chattopadhyay, M. Makino, T. N. Fredrickson, A. W. Hugin, and J. W. Hartley.** 1992. Retrovirus-induced immunodeficiency in the mouse: MAIDS as a model for AIDS. Aids **6:**607-621.
155. **Mosier, D. E., R. A. Yetter, and H. C. Morse, 3rd.** 1987. Functional T lymphocytes are required for a murine retrovirus-induced immunodeficiency disease (MAIDS). J Exp Med **165:**1737-1742.
156. **Muckenfuss, H., M. Hamdorf, U. Held, M. Perkovic, J. Lower, K. Cichutek, E. Flory, G. G. Schumann, and C. Munk.** 2006. APOBEC3 proteins inhibit human LINE-1 retrotransposition. J Biol Chem **281:**22161-22172.
157. **Mulder, L. C., A. Harari, and V. Simon.** 2008. Cytidine deamination induced HIV-1 drug resistance. Proc Natl Acad Sci U S A **105:**5501-5506.
158. **Mullins, J. I., L. Heath, J. P. Hughes, J. Kicha, S. Styrchak, K. G. Wong, U. Rao, A. Hansen, K. S. Harris, J. P. Laurent, D. Li, J. H. Simpson, J. M. Essigmann, L. A. Loeb, and J. Parkins.** 2011. Mutation of HIV-1 genomes in a clinical population treated with the mutagenic nucleoside KP1461. PLoS One **6:**e15135.
159. **Muriaux, D., and J. L. Darlix.** 2010. Properties and functions of the nucleocapsid protein in virus assembly. RNA Biol **7:**744-753.
160. **Neil, S. J., T. Zang, and P. D. Bieniasz.** 2008. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature **451:**425-430.
161. **Newman, E. N., R. K. Holmes, H. M. Craig, K. C. Klein, J. R. Lingappa, M. H. Malim, and A. M. Sheehy.** 2005. Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. Curr Biol **15:**166-170.
162. **Nguyen, D. H., S. Gummuluru, and J. Hu.** 2007. Deamination-independent inhibition of hepatitis B virus reverse transcription by APOBEC3G. J Virol **81:**4465-4472.
163. **Nozaki, A., M. Morimoto, M. Kondo, T. Oshima, K. Numata, S. Fujisawa, T. Kaneko, E. Miyajima, S. Morita, K. Mori, M. Ikeda, N. Kato, and K. Tanaka.** 2010. Hydroxyurea as an inhibitor of hepatitis C virus RNA replication. Arch Virol **155:**601-605.
164. **O'Neil, P. K., G. Sun, H. Yu, Y. Ron, J. P. Dougherty, and B. D. Preston.** 2002. Mutational analysis of HIV-1 long terminal repeats to explore the relative contribution of reverse transcriptase and RNA polymerase II to viral mutagenesis. J Biol Chem **277:**38053-38061.
165. **Ohnota, H., Y. Okada, H. Ushijima, T. Kitamura, K. Komuro, and T. Mizuochi.** 1990. 3'-Azido-3'-deoxythymidine prevents induction of murine

acquired immunodeficiency syndrome in C57BL/10 mice infected with LP-BM5 murine leukemia viruses, a possible animal model for antiretroviral drug screening. Antimicrob Agents Chemother **34:**605-609.

166. **Ostrowski, M. A., T. W. Chun, S. J. Justement, I. Motola, M. A. Spinelli, J. Adelsberger, L. A. Ehler, S. B. Mizell, C. W. Hallahan, and A. S. Fauci.** 1999. Both memory and CD45RA+/CD62L+ naive CD4(+) T cells are infected in human immunodeficiency virus type 1-infected individuals. J Virol **73:**6430-6435.
167. **Pace, C., J. Keller, D. Nolan, I. James, S. Gaudieri, C. Moore, and S. Mallal.** 2006. Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with APOBEC3G and vif genetic variation. J Virol **80:**9259-9269.
168. **Paces, V., J. Doskocil, and F. Sorm.** 1968. Incorporation of 5-azacytidine into nucleic acids of Escherichia coli. Biochim Biophys Acta **161:**352-360.
169. **Pariente, N., A. Airaksinen, and E. Domingo.** 2003. Mutagenesis versus inhibition in the efficiency of extinction of foot-and-mouth disease virus. J Virol **77:**7131-7138.
170. **Pariente, N., S. Sierra, P. R. Lowenstein, and E. Domingo.** 2001. Efficient virus extinction by combinations of a mutagen and antiviral inhibitors. J Virol **75:**9723-9730.
171. **Pathak, V. K., and H. M. Temin.** 1992. 5-Azacytidine and RNA secondary structure increase the retrovirus mutation rate. J Virol **66:**3093-3100.
172. **Pattengale, P. K., C. R. Taylor, P. Twomey, S. Hill, J. Jonasson, T. Beardsley, and M. Haas.** 1982. Immunopathology of B-cell lymphomas induced in C57BL/6 mice by dualtropic murine leukemia virus (MuLV). Am J Pathol **107:**362-377.
173. **Pear, W. S., J. P. Miller, L. Xu, J. C. Pui, B. Soffer, R. C. Quackenbush, A. M. Pendergast, R. Bronson, J. C. Aster, M. L. Scott, and D. Baltimore.** 1998. Efficient and rapid induction of a chronic myelogenous leukemia-like myeloproliferative disease in mice receiving P210 bcr/abl-transduced bone marrow. Blood **92:**3780-3792.
174. **Perales, C., R. Agudo, and E. Domingo.** 2009. Counteracting quasispecies adaptability: extinction of a ribavirin-resistant virus mutant by an alternative mutagenic treatment. PLoS One **4:**e5554.
175. **Perales, C., R. Agudo, S. C. Manrubia, and E. Domingo.** 2011. Influence of mutagenesis and viral load on the sustained low-level replication of an RNA virus. J Mol Biol **407:**60-78.
176. **Perales, C., R. Agudo, H. Tejero, S. C. Manrubia, and E. Domingo.** 2009. Potential benefits of sequential inhibitor-mutagen treatments of RNA virus infections. PLoS Pathog **5:**e1000658.
177. **Perelson, A. S., A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho.** 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. Science **271:**1582-1586.

178. **Peyret, N., P. A. Seneviratne, H. T. Allawi, and J. SantaLucia, Jr.** 1999. Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. Biochemistry **38:**3468-3477.
179. **Pfeiffer, J. K., and K. Kirkegaard.** 2003. A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. Proc Natl Acad Sci U S A **100:**7289-7294.
180. **Phuphuakrat, A., R. Kraiwong, C. Boonarkart, D. Lauhakirti, T. H. Lee, and P. Auewarakul.** 2008. Double-stranded RNA adenosine deaminases enhance expression of human immunodeficiency virus type 1 proteins. J Virol **82:**10864-10872.
181. **Rabson, A. B., and M. A. Martin.** 1985. Molecular organization of the AIDS retrovirus. Cell **40:**477-480.
182. **Rausch, J. W., L. Chelico, M. F. Goodman, and S. F. Le Grice.** 2009. Dissecting APOBEC3G substrate specificity by nucleoside analog interference. J Biol Chem **284:**7047-7058.
183. **Reagan-Shaw, S., M. Nihal, and N. Ahmad.** 2008. Dose translation from animal to human studies revisited. Faseb J **22:**659-661.
184. **Refsland, E. W., M. D. Stenglein, K. Shindo, J. S. Albin, W. L. Brown, and R. S. Harris.** 2010. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. Nucleic Acids Res **38:**4274-4284.
185. **Rogozin, I. B., V. N. Babenko, L. Milanesi, and Y. I. Pavlov.** 2003. Computational analysis of mutation spectra. Brief Bioinform **4:**210-227.
186. **Rogozin, I. B., and Y. I. Pavlov.** 2003. Theoretical analysis of mutation hotspots and their DNA sequence context specificity. Mutat Res **544:**65-85.
187. **Rogstad, D. K., J. L. Herring, J. A. Theruvathu, A. Burdzy, C. C. Perry, J. W. Neidigh, and L. C. Sowers.** 2009. Chemical decomposition of 5-aza-2'-deoxycytidine (Decitabine): kinetic analyses and identification of products by NMR, HPLC, and mass spectrometry. Chem Res Toxicol **22:**1194-1204.
188. **Rosenberg, B. R., and F. N. Papavasiliou.** 2007. Beyond SHM and CSR: AID and related cytidine deaminases in the host response to viral infection. Adv Immunol **94:**215-244.
189. **Ross, L., M. L. Lim, Q. Liao, B. Wine, A. E. Rodriguez, W. Weinberg, and M. Shaefer.** 2007. Prevalence of antiretroviral drug resistance and resistance-associated mutations in antiretroviral therapy-naive HIV-infected individuals from 40 United States cities. HIV Clin Trials **8:**1-8.
190. **Rossi, L., G. Brandi, A. Fraternale, G. F. Schiavano, L. Chiarantini, and M. Magnani.** 1993. Inhibition of murine retrovirus-induced immunodeficiency disease by dideoxycytidine and dideoxycytidine 5'-triphosphate. J Acquir Immune Defic Syndr **6:**1179-1186.
191. **Rossi, L., S. Serafini, P. Franchetti, A. Casabianca, C. Orlandi, G. F. Schiavano, A. Carnevali, and M. Magnani.** 2002. Inhibition of murine AIDS by a heterodinucleotide of azidothymidine and 9-(R)-2-(phosphonomethoxypropyl)adenine. J Antimicrob Chemother **50:**639-647.

192. **Sadler, H. A., M. D. Stenglein, R. S. Harris, and L. M. Mansky.** 2010. APOBEC3G contributes to HIV-1 variation through sublethal mutagenesis. J Virol **84:**7396-7404.
193. **Sakamoto, H., M. Kitano, Y. Suetomi, Y. Takeyama, H. Ohyanagi, T. Nakai, C. Yasuda, and M. Kudo.** 2006. Comparison of standard-dose and low-dose gemcitabine regimens in pancreatic adenocarcinoma patients: a prospective randomized trial. J Gastroenterol **41:**70-76.
194. **Sanjuan, R., J. M. Cuevas, V. Furio, E. C. Holmes, and A. Moya.** 2007. Selection for robustness in mutagenized RNA viruses. PLoS Genet **3:**e93.
195. **SantaLucia, J., Jr.** 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci U S A **95:**1460-1465.
196. **Sasada, A., A. Takaori-Kondo, K. Shirakawa, M. Kobayashi, A. Abudu, M. Hishizawa, K. Imada, Y. Tanaka, and T. Uchiyama.** 2005. APOBEC3G targets human T-cell leukemia virus type 1. Retrovirology **2:**32.
197. **Scadden, D. T., M. Zeira, A. Woon, Z. Wang, L. Schieve, K. Ikeuchi, B. Lim, and J. E. Groopman.** 1990. Human immunodeficiency virus infection of human bone marrow stromal fibroblasts. Blood **76:**317-322.
198. **Schafer, A., H. P. Bogerd, and B. R. Cullen.** 2004. Specific packaging of APOBEC3G into HIV-1 virions is mediated by the nucleocapsid domain of the gag polyprotein precursor. Virology **328:**163-168.
199. **Schumacher, A. J., G. Hache, D. A. Macduff, W. L. Brown, and R. S. Harris.** 2008. The DNA deaminase activity of human APOBEC3G is required for Ty1, MusD, and human immunodeficiency virus type 1 restriction. J Virol **82:**2652-2660.
200. **Schumacher, A. J., D. V. Nissley, and R. S. Harris.** 2005. APOBEC3G hypermutates genomic DNA and inhibits Ty1 retrotransposition in yeast. Proc Natl Acad Sci U S A **102:**9854-9859.
201. **Senavirathne, G., M. Jaszczur, P. A. Auerbach, T. G. Upton, L. Chelico, M. F. Goodman, and D. Rueda.** 2012. Single-stranded DNA scanning and deamination by APOBEC3G cytidine deaminase at single molecule resolution. J Biol Chem **287:**15826-15835.
202. **Sheehy, A. M., N. C. Gaddis, J. D. Choi, and M. H. Malim.** 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature **418:**646-650.
203. **Sheehy, A. M., N. C. Gaddis, and M. H. Malim.** 2003. The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. Nat Med **9:**1404-1407.
204. **Shindo, K., A. Takaori-Kondo, M. Kobayashi, A. Abudu, K. Fukunaga, and T. Uchiyama.** 2003. The enzymatic activity of CEM15/Apobec-3G is essential for the regulation of the infectivity of HIV-1 virion but not a sole determinant of its antiviral activity. J Biol Chem **278:**44412-44416.
205. **Sierra, M., A. Airaksinen, C. Gonzalez-Lopez, R. Agudo, A. Arias, and E. Domingo.** 2007. Foot-and-mouth disease virus mutant with decreased

sensitivity to ribavirin: implications for error catastrophe. J Virol **81:**2012-2024.

206. **Sierra, S., M. Davila, P. R. Lowenstein, and E. Domingo.** 2000. Response of foot-and-mouth disease virus to increased mutagenesis: influence of viral load and fitness in loss of infectivity. J Virol **74:**8316-8323.
207. **Simard, C., M. Huang, and P. Jolicoeur.** 1995. Establishment of leukemic T-cell lines from mice inoculated with the MAIDS defective virus. Virology **206:**555-563.
208. **Sluis-Cremer, N., N. A. Temiz, and I. Bahar.** 2004. Conformational changes in HIV-1 reverse transcriptase induced by nonnucleoside reverse transcriptase inhibitor binding. Curr HIV Res **2:**323-332.
209. **Smith, R. A., L. A. Loeb, and B. D. Preston.** 2005. Lethal mutagenesis of HIV. Virus Res **107:**215-228.
210. **Smith, R. A., K. M. Remington, R. M. Lloyd, Jr., R. F. Schinazi, and T. W. North.** 1997. A novel Met-to-Thr mutation in the YMDD motif of reverse transcriptase from feline immunodeficiency virus confers resistance to oxathiolane nucleosides. J Virol **71:**2357-2362.
211. **Spence, R. A., W. M. Kati, K. S. Anderson, and K. A. Johnson.** 1995. Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors. Science **267:**988-993.
212. **Stenglein, M. D., M. B. Burns, M. Li, J. Lengyel, and R. S. Harris.** 2010. APOBEC3 proteins mediate the clearance of foreign DNA from human cells. Nat Struct Mol Biol **17:**222-229.
213. **Stenglein, M. D., and R. S. Harris.** 2006. APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. J Biol Chem **281:**16837-16841.
214. **Suspene, R., D. Guetard, M. Henry, P. Sommer, S. Wain-Hobson, and J. P. Vartanian.** 2005. Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases in vitro and in vivo. Proc Natl Acad Sci U S A **102:**8321-8326.
215. **Tachedjian, G., and S. P. Goff.** 2003. The effect of NNRTIs on HIV reverse transcriptase dimerization. Curr Opin Investig Drugs **4:**966-973.
216. **Thielen, B. K., K. C. Klein, L. W. Walker, M. Rieck, J. H. Buckner, G. W. Tomblingson, and J. R. Lingappa.** 2007. T cells contain an RNase-insensitive inhibitor of APOBEC3G deaminase activity. PLoS Pathog **3:**1320-1334.
217. **Thomas, J. A., and R. J. Gorelick.** 2008. Nucleocapsid protein function in early infection processes. Virus Res **134:**39-63.
218. **Turelli, P., B. Mangeat, S. Jost, S. Vianin, and D. Trono.** 2004. Inhibition of hepatitis B virus replication by APOBEC3G. Science **303:**1829.
219. **Turville, S. G., S. Peretti, and M. Pope.** 2006. Lymphocyte-dendritic cell interactions and mucosal acquisition of SIV/HIV infection. Curr Opin HIV AIDS **1:**3-9.
220. **Urisman, A., R. J. Molinaro, N. Fischer, S. J. Plummer, G. Casey, E. A. Klein, K. Malathi, C. Magi-Galluzzi, R. R. Tubbs, D. Ganem, R. H. Silverman, and**

**J. L. DeRisi.** 2006. Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant. PLoS Pathog **2:**e25.

221. **Vacharaksa, A., A. C. Asrani, K. H. Gebhard, C. E. Fasching, R. A. Giacaman, E. N. Janoff, K. F. Ross, and M. C. Herzberg.** 2008. Oral keratinocytes support non-replicative infection and transfer of harbored HIV-1 to permissive cells. Retrovirology **5:**66.
222. **Van Damme, N., D. Goff, C. Katsura, R. L. Jorgenson, R. Mitchell, M. C. Johnson, E. B. Stephens, and J. Guatelli.** 2008. The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. Cell Host Microbe **3:**245-252.
223. **van der Kuyl, A. C., and B. Berkhout.** 2012. The biased nucleotide composition of the HIV genome: a constant factor in a highly variable virus. Retrovirology **9:**92.
224. **Vartanian, J. P., D. Guetard, M. Henry, and S. Wain-Hobson.** 2008. Evidence for editing of human papillomavirus DNA by APOBEC3 in benign and precancerous lesions. Science **320:**230-233.
225. **Vazquez-Perez, J. A., C. E. Ormsby, R. Hernandez-Juan, K. J. Torres, and G. Reyes-Teran.** 2009. APOBEC3G mRNA expression in exposed seronegative and early stage HIV infected individuals decreases with removal of exposure and with disease progression. Retrovirology **6:**23.
226. **Vodicka, M. A., W. C. Goh, L. I. Wu, M. E. Rogel, S. R. Bartz, V. L. Schweickart, C. J. Raport, and M. Emerman.** 1997. Indicator cell lines for detection of primary strains of human and simian immunodeficiency viruses. Virology **233:**193-198.
227. **Watts, J. M., K. K. Dang, R. J. Gorelick, C. W. Leonard, J. W. Bess, Jr., R. Swanstrom, C. L. Burch, and K. M. Weeks.** 2009. Architecture and secondary structure of an entire HIV-1 RNA genome. Nature **460:**711-716.
228. **Wedekind, J. E., G. S. Dance, M. P. Sowden, and H. C. Smith.** 2003. Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. Trends Genet **19:**207-216.
229. **Wong, A., R. A. Soo, W. P. Yong, and F. Innocenti.** 2009. Clinical pharmacology and pharmacogenetics of gemcitabine. Drug Metab Rev **41:**77-88.
230. **Wu, W., L. E. Henderson, T. D. Copeland, R. J. Gorelick, W. J. Bosche, A. Rein, and J. G. Levin.** 1996. Human immunodeficiency virus type 1 nucleocapsid protein reduces reverse transcriptase pausing at a secondary structure near the murine leukemia virus polypurine tract. J Virol **70:**7132-7142.
231. **Xia, Q., J. Radzio, K. S. Anderson, and N. Sluis-Cremer.** 2007. Probing nonnucleoside inhibitor-induced active-site distortion in HIV-1 reverse transcriptase by transient kinetic analyses. Protein Sci **16:**1728-1737.
232. **Yetter, R. A., R. M. Buller, J. S. Lee, K. L. Elkins, D. E. Mosier, T. N. Fredrickson, and H. C. Morse, 3rd.** 1988. CD4+ T cells are required for

development of a murine retrovirus-induced immunodeficiency syndrome (MAIDS). J Exp Med **168:**623-635.

233. **Yoo, C. B., and P. A. Jones.** 2006. Epigenetic therapy of cancer: past, present and future. Nat Rev Drug Discov **5:**37-50.

234. **Yu, Q., R. Konig, S. Pillai, K. Chiles, M. Kearney, S. Palmer, D. Richman, J. M. Coffin, and N. R. Landau.** 2004. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. Nat Struct Mol Biol **11:**435-442.

235. **Zeng, M., P. J. Southern, C. S. Reilly, G. J. Beilman, J. G. Chipman, T. W. Schacker, and A. T. Haase.** 2012. Lymphoid tissue damage in HIV-1 infection depletes naive T cells and limits T cell reconstitution after antiretroviral therapy. PLoS Pathog **8:**e1002437.

236. **Zennou, V., and P. D. Bieniasz.** 2006. Comparative analysis of the antiretroviral activity of APOBEC3G and APOBEC3F from primates. Virology **349:**31-40.

237. **Zennou, V., D. Perez-Caballero, H. Gottlinger, and P. D. Bieniasz.** 2004. APOBEC3G incorporation into human immunodeficiency virus type 1 particles. J Virol **78:**12058-12061.

238. **Zhang, H., B. Yang, R. J. Pomerantz, C. Zhang, S. C. Arunachalam, and L. Gao.** 2003. The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. Nature **424:**94-98.

239. **Zhang, J.** 2004. Host RNA polymerase II makes minimal contributions to retroviral frame-shift mutations. J Gen Virol **85:**2389-2395.

240. **Zheng, Y. H., D. Irwin, T. Kurosu, K. Tokunaga, T. Sata, and B. M. Peterlin.** 2004. Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. J Virol **78:**6073-6076.

241. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res **31:**3406-3415.

# APPENDIX I

## DECLARATION OF CONTRIBUTIONS TO CO-AUTHORED PUBLICATION: ANALYSIS OF THE EX VIVO AND IN VIVO ANTIRETROVIRAL ACTIVITY OF GEMCITABINE

**I am a co-author on the following publication:**

Clouser, C.L., **Holtz, C.M**., Mullet, M., Crankshaw, D.L., Briggs, J.E., Chauhan, J., Patterson, S.E., and L.M. Mansky. 2011. Analysis of the ex vivo and in vivo antiretroviral activity of gemcitabine. *PLoS One* 6:e15840

My contributions to this publication are as follows:

- Generated data in Figure A1-1A
- Contributed text for data generated and edited a complete draft of the manuscript

## Introduction

Retroviruses are a significant source of morbidity and mortality worldwide. For example, there are 33 million people infected with human immunodeficiency virus type 1 (HIV-1) whereas the retrovirus, xenotropic murine leukemia virus like-virus (XMRV), has recently been proposed to be linked to prostate cancer and chronic fatigue syndrome (CFS) (133, 220). Although there are a number of drugs available for HIV-1 chemotherapy, the efficacy of these treatments is limited by the emergence of drug resistance, cost of treatment, and off-target effects. These limitations necessitate the development of new drugs and novel drug targets for HIV as well as other retroviruses. Similarly, if XMRV is shown to be the etiological agent of either prostate cancer and/or CFS, the development of new drugs could reduce morbidity and mortality.

Current anti-retroviral drugs target viral proteins that are necessary for viral replication and production. Under suboptimal therapy, the combination of both the high rates of replication and mutation leads to the emergence of drug resistance. Although drugs that target host proteins could delay or prevent the emergence of drug resistance, there are significant side effects associated with inhibiting host proteins. While cellular deoxynucleoside triphosphates (dNTPs) are necessary for host cell function, even small changes in dNTP pools appear to affect viral replication without significant cellular toxicity (17).

Previous studies have shown that hydroxyurea, which inhibits the cellular enzyme ribonucleotide reductase effectively decreases replication of HIV-1 and was recently shown to also inhibit hepatitis C virus replication as well (135, 146, 163). Nonetheless, several factors make hydroxyurea undesirable as an antiviral including 1)

pharmacokinetics of hydroxyurea vary from person to person making its plasma levels unpredictable and 2) significant toxicities (pancreatitis, hepatotoxicity) are associated with the use of hydroxyurea when used to treat HIV-1 infection. The toxicity of hydroxyurea is further emphasized in the murine AIDS (MAIDS) model in which all animals treated with hydroxyurea died from drug-related toxicities (146).

Alternatives to hydroxyurea that have a more desirable pharmacokinetic profile and lower toxicity concerns may offer a new and useful treatment for retroviral infections. In fact, other ribonucleotide reductase inhibitors have been shown to be more effective and less toxic in the MAIDS model (145). However, no follow up studies have been published regarding their potential for clinical use.

2′,2′-Difluoro-2′-deoxycytidine, commonly referred to as gemcitabine represents a potential alternative to hydroxyurea as it has been shown to have two mechanisms of action including inhibition of ribonucleotide reductase (98, 229). Gemcitabine is clinically-approved for cancer therapy and its anti-cancer mechanism is attributed to its ability to inhibit ribonucleotide reductase thereby limiting dNTP pools available for DNA synthesis in cancer cells. Since retroviruses may be more sensitive to dNTP pool alterations than cellular polymerases, we hypothesized that gemcitabine would be an alternative to hydroxyurea that could be translated to clinical use for the treatment of retroviral infections when used in combination with current anti-retroviral therapies.

In a recent study, we identified two clinically-approved drugs, decitabine and gemcitabine, that had potent anti-HIV activity in cell culture. The potency by which gemcitabine inhibited infectivity suggested that it may be useful for the treatment of retroviral infections when used with decitabine or in combination with current

antiretroviral therapies.

In this study, we examined the ability of gemcitabine to inhibit replication of another retrovirus, murine leukemia virus in cell culture. Additionally, we examined the efficacy and toxicity of gemcitabine in vivo using LP-BM5 murine leukemia virus (LP-BM5 MuLV, a murine AIDS model). We chose this animal model as it has been used extensively to screen potential anti-HIV drugs and has been validated with a number of clinically approved anti-HIV drugs such as AZT and PMPA (tenofovir) that have relatively broad antiretroviral activity (67, 165, 191). In this study, we show that gemcitabine decreased infectivity of MuLV in cell culture with no detectable cytotoxicity. Similarly, gemcitabine decreased disease progression in the MAIDS model at non-toxic doses although toxicities were detected at doses just three times that of the effective dose. Our findings, along with previous observations (50), indicate that gemcitabine has relatively broad antiretroviral activity with minimal toxicity and could be useful for in vivo antiretroviral combination therapy.

## Results

*Gemcitabine inhibits MuLV in cell culture*

Before examining the antiviral activity of gemcitabine *in vivo*, we first examined the ability of gemcitabine to inhibit MuLV in cell culture. To do this, a GFP-tagged MuLV was pseudotyped with VSV-G and used to infect target cells that had been pretreated with increasing concentrations of gemcitabine. Flow cytometry was then used to determine the percentage of infected cells. As shown in Fig. A1-1A, gemcitabine potently decreased MuLV infectivity in a concentration-dependent manner with an EC50

in the low nM range. Additionally, there was no toxicity seen at the concentrations of gemcitabine needed to inhibit viral replication when the cells were exposed to gemcitabine for the same time as was used to assess gemcitabine's effect on infectivity (Fig. A1-1B).

*Gemcitabine inhibits progression of MAIDS as detected by spleen weight and histopathology*

The data from Fig. A1-1 indicate that gemcitabine inhibits replication of MuLV in cell culture. To examine the ability of gemcitabine to inhibit MuLV in an in vivo system, we treated mice infected with LP-BM5 with increasing doses of gemcitabine (Table A1-1). Progression of MAIDS was characterized by extensive lymphoproliferation, splenomegaly, increased IgM levels, the development of lymphoma, and increased susceptibility to infection (155, 172, 207, 232).

Proliferation of lymphoid cells in the spleen contributes to the splenomegaly observed in mice infected with LP-BM5 and is an indicator of disease progression and disease severity (232). To determine if gemcitabine decreases splenomegaly, spleens were obtained and weighed at the time of sacrifice. As expected, the ratio of spleen to body weight was significantly increased in infected animals that did not receive treatment (Fig. A1-2). Mice treated with 1 or 2 mg/kg/day had an average ratio of spleen to body weight that was significantly lower than the infected, untreated mice. In fact, there was no significant difference in the spleen to body weight ratio between the treated mice and those that were not infected (Fig. A1-2).

Since gemcitabine significantly decreased spleen size in mice infected with LP-

BM5, we next examined whether the decrease in spleen size correlated with a decrease in the histopathological changes characteristic of MAIDS. Table A1-2 demonstrates that there were no significant findings in spleen from two of the four uninfected mice. The spleens of the other two uninfected mice were identified as either reactive or as having a score of 1 (see Materials and Methods for scoring system). Of all other groups, the infected, untreated mice had the most severe lesions as the spleens from all four of the mice in this group received a score of 2 or higher, indicating extensive changes in splenic architecture and significant expansion of lymphoid cells. In general, there was a decrease in splenic lesion score as the dose of administered gemcitabine increased. For example, all spleens from mice receiving the highest dose of gemcitabine were normal, while 5 of the 6 mice in the 3 mg/kg/day treatment group had spleens with no significant findings and 4 of the 6 mice in the 2 mg/kg/day treatment group were within normal limits. Finally, 1 of the 7 mice from the 1 mg/kg/day treatment group had a spleen with normal histology while the other 6 spleens received a score of 1, indicating mild lesions.

*Treatment with gemcitabine decreases MAIDS-associated lesions in the lymph nodes*

Infection with LP-BM5 destroys the architecture of lymph nodes and alters the composition (232). When detected, the lymph nodes from infected or uninfected mice were examined by histopathology as described in the Materials and Methods section. Lymph nodes were only detectable in one of the four uninfected mice (Table A1-3). These lymph nodes were scored as a 1, indicating an enlargement of the lymph nodes with diffuse sheets of medium to large lymphoid cells while maintaining corticomedullary architecture. Consistent with the spleen data, the size and lesions of the

lymph nodes was greatest in mice that were infected but not treated (Table A1-3). Additionally, lymph nodes became more difficult to detect as the dose of gemcitabine increased, indicating that gemcitabine decreased the extent of lymph node enlargement. Consistent with this, the pathology of lymph nodes from mice treated with the higher doses of gemcitabine had lower scores indicating less severe lesions. For example, 6 of the 7 mice treated with 1 mg/kg/day of gemcitabine received a score of 1, whereas all of the lymph nodes from mice treated with either 2 or 3 mg/kg/day received scores that were consistent with a decrease in disease progression when compared to the untreated but infected mice or mice treated with 1 mg/kg/day of gemcitabine.

*Effect of gemcitabine on plasma IgM levels*

Soon after infection with LP-BM5, mice demonstrate a significant increase in plasma IgM levels that peak approximately 8–9 weeks post-infection (172, 232). Subsequently, IgM levels decrease, but remain elevated compared to uninfected animals. To examine the effect of gemcitabine on IgM levels, plasma was isolated from whole blood collected from mice at the time of sacrifice. As expected, infected but untreated animals demonstrated a significant increase in IgM levels compared to the uninfected animals (Fig. A1-3). In contrast, treatment with 1 or 2 mg/kg/day of gemcitabine significantly decreased IgM levels compared to the untreated animals. Surprisingly, IgM levels from mice treated with 2 mg/kg/day of gemcitabine were significantly lower than that seen in the uninfected animals.

*Gemcitabine decreases provirus levels*

The data pathology indicated that gemcitabine decreases the severity or progression of murine AIDS. However, to more directly determine if gemcitabine inhibits replication of LP-BM5, provirus levels were quantified from spleen obtained at the time of sacrifice. Levels of the defective provirus were normalized to 18S rRNA levels as previously described (39). The data shown in Fig. A1-4 revealed high levels of provirus in mice that were infected, but untreated. Mice treated with either 1 or 2 mg/kg/day of gemcitabine had negligible provirus levels that were comparable to the uninfected control mice, indicating reduced viral replication.

*Toxicity of gemcitabine as determined by changes in body weight*

Body weight is often used as an indicator of toxicity in laboratory animals. In this study, animals losing 15% or more of their body mass were euthanized and such loss of weight was attributed to drug-related toxicity. All mice treated with 3 and 4 mg/kg/day of gemcitabine, regardless of infection status, lost 15% or more body mass and were euthanized before the end of the study. Fig. A1-5 shows the change in body weight of all mice at the time of euthanasia. The data demonstrates that mice treated with 2 mg/kg/day of gemcitabine or higher lost body weight while untreated mice or mice treated with 1 mg/kg/day of gemcitabine gained body weight throughout the study.

**Discussion**

In the absence of a cure or vaccine for HIV-1 infection, the identification of novel drug targets and the development of new drugs is the best approach to address the emergence of resistance as well as the complications associated with current therapies.

Drugs that target cellular proteins are less likely to be susceptible to the emergence of drug resistance compared to the current anti-HIV therapies which target viral proteins. However, drugs that target cellular proteins are likely to be associated with an increase in toxicity which may limit their clinical use. For example, hydroxyurea has been used alone and in combination with nucleoside reverse transcriptase inhibitors (NRTIs) to decrease viral loads in HIV-1 infected individuals. The side effects associated with hydroxyurea has significantly curtailed its clinical use. Less toxic alternatives to hydroxyurea would offer important alternatives for the treatment of HIV-1 as well as other retroviral infections.

A recent study demonstrated that two clinically approved drugs, decitabine and gemcitabine have potent anti-HIV activity in cell culture when used alone or in combination (50). Importantly, based on gemcitabine's potency and lack of toxicity, the data suggest that gemcitabine may be a clinically relevant alternative to hydroxyurea for the treatment of retroviral infections.

In this study, we examined the antiretroviral activity of gemcitabine *in vivo* and *ex vivo* using the LP-BM5 MuLV model (a murine AIDS model). Murine AIDS is caused by a combination of three murine leukemia viruses including BM5-eco, BM5def, and mink cell focus inducing virus (MCF) (42, 43). BM5def is a retrovirus that is unable to replicate due to deletion of most of the pol gene. However, its intact gag gene is thought to be responsible for MAIDS pathogenesis. Although LP-BM5 is not a perfect model for AIDS pathogenesis, it is well characterized, safe, inexpensive, and has been validated with a number of clinically approved anti-HIV-1 drugs and has been useful for identifying drugs with broad antiretroviral activity (for review see (67)).

Our studies found that gemcitabine decreased replication of MuLV in cell culture as well as in vivo using the MAIDS model. In fact, all measures of MAIDS-associated pathogenesis including splenomegaly, histopathology of spleen and lymph nodes and levels of IgM and provirus indicated reduced viral replication. The reduction of provirus by gemcitabine further supported the gemcitabine-mediated reduction in viral replication. There are at least two mechanisms that could account for the decrease in disease progression. First, gemcitabine could decrease dNTP levels enough to decrease proliferation of infected cells. Second, gemcitabine could alter dNTP levels such that viral replication is inhibited without inhibition of cellular proliferation. These mechanisms are likely to be dose dependent with low gemcitabine doses preferentially inhibiting viral replication and higher doses inhibiting both viral replication and cell proliferation. In support of this, 2 mg/kg/day of gemcitabine caused a significant reduction in IgM levels compared to the uninfected animals. This is likely due to gemcitabine-mediated decrease in B cell proliferation. In contrast, IgM levels from mice treated with 1 mg/kg/day had levels comparable to the uninfected animals which suggest there was not a significant reduction in B cell numbers compared to the uninfected mice. The cell culture data also supported this assertion as the decrease in replication was not associated with a significant loss of cell numbers (Fig. A1-1B). These data indicate that gemcitabine's antiviral activity is distinct from its activity used to treat human cancers. Furthermore, doses of gemcitabine used clinically in cancer treatment are significantly higher than those used in our study. For example, gemcitabine's anti-cancer activity is achieved with a dosing regimen that includes 1000 mg/m$^2$ given once every week for seven weeks, followed by one week without drug and additional rounds of treatment as

needed (193). In our study, an antiviral effect was seen in animals at doses as low as 1 mg/kg which correlates to 3 mg/$m^2$ when using the BSA method to convert the mouse dose to the human equivalent dose (mouse dose multiplied by mouse km of 3) (183). Although the metabolism of gemcitabine in mice is likely different compared to humans, the significant difference in dosing further supports that the antiviral activity seen here is not due to inhibition of cell proliferation.

Toxicity of gemcitabine was significant at higher doses. In fact, all mice treated with 3 and 4 mg/kg/day lost 15% or greater of their body weight during the study and as a result were prematurely euthanized. However, the 1mg/kg/day dose of gemcitabine decreased disease pathology with no detectable toxicity. Specifically, all of these animals gained body weight similar to the uninfected animals. Additionally, these animals did not show any signs of hepatotoxicity as detected by histopathological analysis (data not shown). The 2 mg/kg/day dose was also effective at decreasing disease progression. However, some of the mice in this treatment group showed signs of toxicity in the loss of body weight while others in this group appeared to tolerate gemcitabine well. Like the animals treated with 1 mg/kg/day gemcitabine, there was no hepatotoxicity observed in animals treated with 2 mg/kg/day gemcitabine. The extent to which gemcitabine effected the animals in the 2 mg/kg group was not limited to toxicity as the efficacy of gemcitabine's antiviral activity also varied (see Figure A1-2). Although it is not clear why the mice could have such different responses to the drug, the affect of this variation could be minimized by increasing the number of animals in each treatment group.

Although gemcitabine has been used as cancer chemotherapy for many years, it has likely been neglected as an antiviral due to its poor oral bioavailability, which would

necessitate drug delivery by injection. However, a prodrug for gemcitabine has been described recently, making gemcitabine a more attractive candidate for use in treatment of retroviral infections, including highly drug-resistant HIV-1 (18). If its bioavailability is improved, gemcitabine could serve as a novel anti-HIV drug for those resistant to the current therapies.

In summary, the findings of our study indicate that gemcitabine has potent antiretroviral activity *in vivo* and *ex vivo* using the LP-BM5 MuLV model. These findings, along with previous *ex vivo* HIV-1 studies with gemicitabine and decitabine, suggest that gemcitabine has broad antiretroviral activity and could be particularly useful *in vivo* when used in combination drug therapy.

## Materials and Methods

### *Materials*

SC-1/MuLV LP-BM5 cells, chronically infected with LP-BM5, were obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH from Dr. Herbert Morse and Dr. Janet Hartley (42, 97, 154). 293T cells were obtained from American Tissue Type Culture (ATCC). C57Bl/6 mice were purchased from Jackson Laboratories. DMEM was purchased from MediaGrow. Gemcitabine was obtained from Carbosynth (Berkshire, U.K.). The IgM Elisa kit was from Assay Designs (Ann Arbor, MI). The plasmids, pCR-DEF and pCR-18S were a kind gift from Dr. Mauro Magnani (University of Urbino, Urbino, Italy) and have previously been described (39). U373-MAGI-CXCR4CEM cells were obtained from Dr. Michael Emerman through the AIDS Research and Reference Reagent Program, Division

of AIDS, NIAID, NIH (93, 226). The plasmid, pIRES2-EGFP was obtained from Clontech (Mountain View, CA). The plasmids pMIGR1, pJK3, pL-VSV-G, and CMV-Tat were kind gifts from Vineet Kewal Ramani (NCI-Fredrick). The plasmid pMIGR1 is an MLV vector containing an IRES-GFP element (173); pJK3 contains GagPol driven off HIV-1 LTR (15). pL-VSV-G is also driven off the HIV-1 LTR (15); and CMV-Tat allows transcription off the HIV-1 LTR (15).

*Cell culture*

SC-1/MuLV LP-BM5 cells and 293T cells were maintained in DMEM containing 10% fetal clone 3 (FC3) serum (Hyclone, Logan, VT) and pen/strep at 37°C in 5% CO2. U373-MAGI-CXCR4CEM cells expressing the CD4 receptor and the CXCR4 co-receptor were maintained at 37°C in 5% CO2 in selection media composed of DMEM with 10% FC3, 1 μg/mL puromycin, 0.1 mg/mL hygromycin, and 0.2 mg/mL neomycin.

*Transfection of 293T cells*

pMIGR1 (13.9 μg), pJK3 (6.9 μg), pL-VSV-G (5.54 μg) and CMV-Tat (1 μg) were transfected into 293T cells using the calcium phosphate method. Forty-eight hours after transfection, cell culture supernatant was collected and frozen at −80°C.

*Drug Treatment/Flow Cytometry*

U373-MAGI-CXCR4$_{CEM}$ cells (62,000) were plated in 12-well dishes 24 h prior to drug treatment. Twenty-four hours later, gemcitabine was added to the cultures at the concentrations indicated in the legend to Fig. A1-1. Two hours post-drug treatment, 500

µl of virus was added to each culture. Twenty-four hours later, the medium containing gemcitabine (or DMSO) was removed replaced with new media. Twenty-four hours after the media change, the U373-MAGI-CXCR4$_{CEM}$ cells were collected and analyzed by flow cytometry to determine the percentage of cells expressing GFP. Infectivity was normalized for each individual experiment by setting the infectivity of the untreated cells to 100 for each experiment and then multiplying the data from the other individual treatments by the number used to convert the no drug treated cells to 100.

*Cellular Proliferation*

Cellular proliferation was examined using the CellTiter-Glo kit from Promega according to the manufacturer's instructions. U373-MAGI-CXCR4 cells (4,500/well) were plated in a 96-well dish 24 h prior to drug treatment. Cells were treated (or not) with gemcitabine for 24 h before proliferation was assessed by luciferase activity. DMSO was used as a control for the untreated cells. Ethanol (20%) was used as a positive control for cellular toxicity. The data were converted to relative cell number by setting the no drug treated cells at 100 for each experiment and then multiplying each other sample data by the number used to convert the no drug treated cells to 100. This conversion was normalized for differences in luciferase activity among different experiments.

*Mice*

Female C57BL/6 mice aged 8–10 wk old were purchased from Jackson Laboratories (Sacramento, California) and were housed in standard rodent shoebox caging without a filter top at 22±1°C with a 12 h light/dark cycle, 60±5% humidity, and

12 air changes/h. Mice were fed lab chow and water *ad libitum*. The experimental protocol was approved by the Department of Veterans Affairs Medical Center (Minneapolis, MN) IACUC committee (IACUC protocol number 0803A28341).

*Infection of mice with LP-BM5 MuLV*

LP-BM5 was produced from confluent SC-1 cells by filtering the cell supernatant through a 0.25 μm syringe. The filtered viral supernatant was maintained at −80°C until inoculation was performed. C57BL/6 mice were inoculated with two intraperitoneal injections of 0.25 mL of viral supernatant or DMEM spaced three days apart.

*Treatment of LP-BM5 MuLV infected mice with gemcitabine*

C57BL/6 mice were randomly divided into 7 groups including: 1) uninfected untreated; 2) uninfected treated with 4 mg/kg/day gemcitabine; 3) infected, untreated; 4) infected treated with 1 mg/kg/day; 5) infected with 2 mg/kg/day; 6) infected, treated with 3 mg/kg/day; and 7) infected, treated with 4 mg/kg/day. All mice were treated daily with either gemcitabine or phosphate buffered saline (PBS) for 8 wk beginning 1 wk post-infection. Animals were weighed daily to achieve proper dosing and to detect changes in body mass due to toxicity or infection. Animals did not receive drug treatment 24 h prior to sacrifice.

*Sacrifice of animals*

Animals were euthanized either 8 wk post-drug treatment or when animals lost 15% or greater body mass due to toxicity associated with treatment. Animals were

weighed and anesthetized with 100 mg/kg ketamine plus 10 mg/kg xylazine prior to blood collection from the submandibular vein. After euthanasia, spleens were removed, weighed, and sectioned for histopathological analysis as well as for the quantification of proviral levels. Lymph nodes, when detected, were obtained and sectioned for histopathological analysis as was liver. Necropsy was performed to assess any gross abnormalities in any organs.

*Splenic histopathology*

Histological examination and scoring of spleen was performed by A.C.V.P. board certified veterinary pathologists from the University of Minnesota Masonic Cancer Center Comparative Pathology Shared Resource. The severity of lymphoproliferation within the spleen was scored based on a previously published scoring system (97), with slight modifications as follows: 1) Normal spleen (N) demonstrated normal architecture and a ratio of while pulp to red pulp (W:R) of ≤ ~1:1. 2) Spleens designated as reactive (R) were defined as having reactive germinal center expansion with a W:R ratio of ~2:1. Germinal centers in these spleens were prominent and expanded. 3) Spleens designated as a score of 1 were characterized by more extensive germinal center expansion that those in the reactive group, with a W:R ratio of ~3:1. Lymphoid nodules in these spleens were not coalescent and had prominent mantle zones. 4) Spleens designated with a score of 2 contained multifocal large lymphoid nodules that coalesced and resulted in effacement of red pulp and mantle zones. The W:R ratio in these spleens was ~4:1. 5) Spleens designated as a score of 3 were characterized by extensive proliferation of medium to large lymphoid cells that nearly completely effaced the red pulp and mantle zones of the

lymphoid nodules. The W:R ratio in these spleens was >4:1. 6) Finally, spleens defined as D were characterized by lymphoid depletion within white pulp.

*Lymph node histopathology*

A scoring system for lymph nodes in MAIDS mice has not been previously reported. Scoring of lymph node lesions was similar to that described for the spleen and was done using following grading scheme: 1) Normal (N) was defined as lymph nodes demonstrating normal architecture. 2) Reactive (R) was defined as lymph nodes in which the corticomedullary architecture was maintained, but the lymphoid follicles had enlarged germinal centers 3) A score of 1 was defined by moderately enlarged lymph nodes with diffuse sheets of medium to large lymphoid cells while still maintaining some follicular structures and corticomedullary architecture. 4) A score of 2 was defined as diffusely enlarged lymph node where the corticomedullary architecture was absent, but a rim of small lymphocytes was present beneath the capsule. Finally, 5) lymph nodes designated with a score of 3 were diffusely enlarged, and contained medium to large lymphocytes that infiltrated the lymph node capsule and extended into adjacent tissues.

*Serum immunoglobulin M determination*

After 8 wk of treatment, IgM levels were determined by ELISA as per the manufacturers' instructions (Assay Designs, Ann Arbor, MI). Plasma was isolated from whole blood by centrifugation at 1,000×g for 10 minutes at room temperature and stored at −80°C until the assay was performed. A standard curve was generated using a standard IgM of known concentration and a correlation coefficient (R2) of >0.98 was considered

acceptable data for analysis. All samples were run in triplicate and to confirm reproducibility between assays, a subset of samples was run in three independent experiments.

*Determination of provirus levels from spleen*

At 8 wk post-treatment, BM5d DNA content was assayed from genomic DNA isolated from spleen. Total cellular DNA was isolated using the Roche kit as per the manufacturers' instructions. BM5d DNA was quantified by real-time PCR assay as previously described (39). The amount of BM5d DNA in spleen was calculated by interpolation of the experimentally determined plasmid standard curve and was normalized to 18S rRNA. SYBR green mastermix (Applied Biosystems) containing genomic DNA (1 μl) and 25 pM of primers. The primers used to detect BM5def were 5′ CCTTTATCGACACTTCCCTTTT 3′ and 5′ TGGCAGAGGAGGAAGGTT3′. The primers used to detects 18S rRNA were: 5′ GTAACCCGTTGAACCCCATT (forward) and 5′CCATCCAATCGGTAGTAGGG (reverse). Conditions for amplification of BM5Def and 18S rRNA included an initial heat activation of the polymerase at 95°C for 13 min followed by 40 cycles of 95°C for 15 sec and 62°C for one min. Samples were then heated to 72°C for 5 min, 95°C for 1 min prior to performing a melt curve analysis from 55°C to 95°C. Data was used for analysis only when the standard curve for each primer set yielded an R2 of >0.99 with an efficiency of 90–110%, single peaks were observed in the melt curves and template controls gave no detectable amplification. Proviral levels normalized to 18S rRNA levels by the Pfaffel method of quantification and were determined in triplicate for each set of reactions and three independent assays

were performed with each sample.

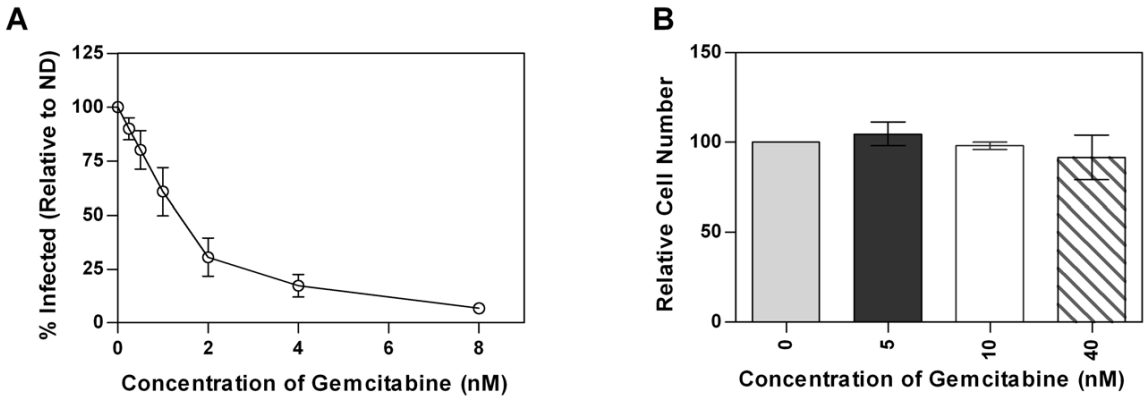



**Figure A1-1. Gemcitabine inhibits MuLV replication in cell culture in the absence of toxicity.** A. Infectivity of MuLV. MuLV containing GFP were produced from 293T cells and used to infect U373-MAGI-CXCR4CEM cells that were treated with the indicated concentrations of gemcitabine. The data represents the average 6 SD of three independent experiments. B. Toxicity of gemcitabine in U373-MAGI-CXCR4CEM cells treated with the indicated concentrations of gemcitabine. The data represents the average 6 SD of three independent experiments.

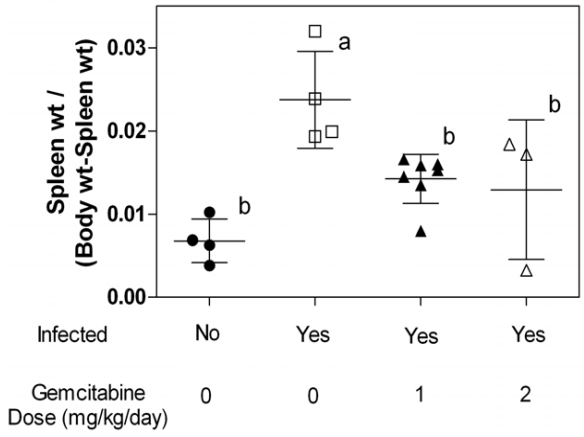



**Figure A1-2. Ratio of spleen weights to body from mice infected with LP-BM5 MuLV.** Each symbol (circles, squares, and triangles) represents one mouse. The average 6 SD is shown. Treatment groups labeled with the same letter (eg. the 1 and 2 mg/kg/day

groups are both labeled with ‘‘b’’) are not statistically different from one another whereas treatment groups labeled with different letters (eg. ‘‘a’’ from one group and ‘‘b’’ for another group) are statistically different from one another as determined by One-Way ANOVA with Tukey-Kramer post-test $p<0.05$. n=4 for the untreated groups, n=7 for the mice treated with 1 mg/kg/day and n = 3 for mice treated with 2 mg/kg/day.

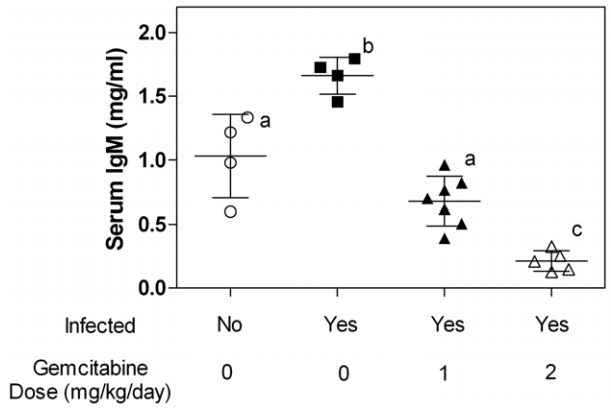



**Figure A1-3. Serum IgM levels from mice infected with LP-BM5 MuLV.** Each symbol (circles, squares and triangles) represent one animal. Treatment groups labeled with different letters are not statistically different from one another while treatment groups labeled with different letters are statistically different as determined by One- Way ANOVA with Tukey-Kramer post-test $p<0.05$. n = 4 for both the untreated groups, n = 7 for animals treated with 1 mg/kg/day, and n = 5 for animals treated with 2 mg/kg/day gemcitabine.

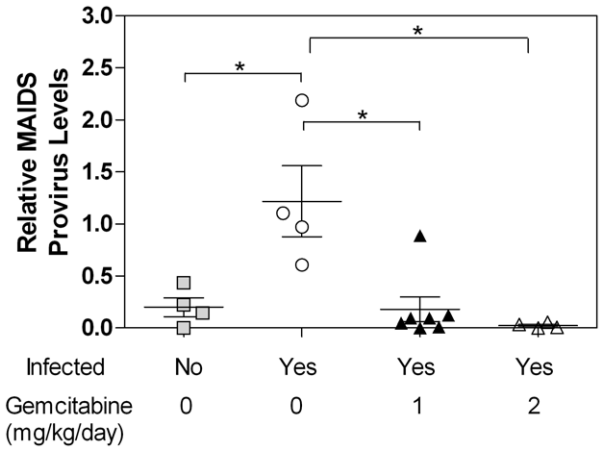



**Figure A1-4. Provirus levels in mice infected with LP-BM5 MuLV.** Genomic DNA was extracted from mouse spleen and quantitative real time PCR was performed to detect the defective MAIDS provirus. Provirus levels were normalized to the 18S rRNA gene. Each symbol (squares, circles, and triangles) represents one animal. The Pfaffel modification of the DDCt method was used to assess gene expression. Statistical significance was assessed by One-Way ANOVA with Tukey-Kramer post-test with $p<0.05$ designated as significant.

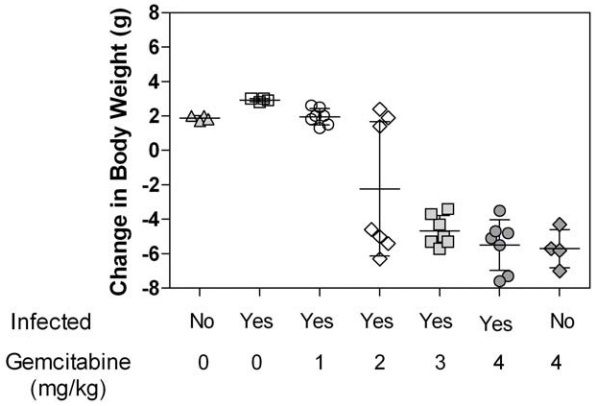



**Figure A-5. Change in body weights of mice infected with LP-BM5 MuLV and treated with the indicated doses of gemcitabine.** Each symbol (triangles, squares,

circles, diamonds, and hatch marks) represent one mouse. The average 6 SD is shown for each treatment group.

**Table A1-1.** Treatment groups for ex vivo analysis of gemcitabine*.

| Treatment Group | Infection Status | Treatment | Number in group at start of study | Number of surviving mice at end of study |
|---|---|---|---|---|
| 1 | Not infected | Saline | 4 | 4 |
| 2 | Infected | Saline | 4 | 4 |
| 3 | Infected | 1 mg/kg/day gemcitabine | 7 | 7 |
| 4 | Infected | 2 mg/kg/day gemcitabine | 7 | 4 |
| 5 | Infected | 3 mg/kg/day gemcitabine | 7 | 0 |
| 6 | Infected | 4 mg/kg/day gemcitabine | 7 | 0 |
| 7 | Not Infected | 4 mg/kg/day gemcitabine | 4 | 0 |

*Treatment Groups, number of animals at the start of the study and the number of animals surviving at the end of the study.

**Table A1-2.** Histopathological analysis of spleen from gemcitabine-treated animals with LP-BM5.*

| Not infected | Infected, untreated | 1 mg/kg/day gemcitabine | 2 mg/kg/day gemcitabine | 3 mg/kg/day gemcitabine | 4 mg/kg/day gemcitabine |
|---|---|---|---|---|---|
| NSF | 2 | 1 | NSF | NSF | NSF |
| NSF | 2 | 1 | NSF | NSF | NSF |
| R | 2 | 1 | Ab | NSF | NSF |
| 1 | 3 | NSF | NSF | NSF | NC |
| | | 1 | NSF | NSF | NC |
| | | 1 | D | Ab | NC |
| | | 1 | NC | NC | NC |

*Sections of spleen were analyzed as described in Materials and Methods. NSF = no significant findings; R = reactive, Ab = Abnormal composition; D = depleted white pulp. NC = tissue was not collected. Score of 3 = high pathology; score of 2 = intermediate pathology; score of 1 = low degree of pathology. Each box corresponds to a different animal.

**Table A1-3.** Histopathological analysis of lymph nodes from gemcitabine-treated animals with LP-BM5 MuLV.*

| Not infected | Infected, untreated | 1 mg/kg/day gemcitabine | 2 mg/kg/day gemcitabine | 3 mg/kg/day gemcitabine | 4 mg/kg/day gemcitabine |
|---|---|---|---|---|---|
| ND | 3 | NSF | NSF | NSF | N/A |
| ND | 3 | NSF | NSF | NSF | N/A |
| ND | 3 | 1 | ND | NSF | N/A |
| 1 | 2 | 1 | ND | N/A | N/A |
| | | 1 | ND | N/A | N/A |
| | | 1 | N/A | ND | N/A |
| | | 2 | N/A | ND | N/A |

*Sections of lymph nodes were analyzed as described in Materials and Methods. NSF = no significant findings; N/A = not applicable because tissues were not collected. ND = lymph nodes from these animals were not detected. Scores of 3 = high degree of pathology; score of 2 = intermediate degree of pathology; score of 1 = low degree of pathology. Each box corresponds to a different animal.

## APPENDIX II

### DECLARATION OF CONTRIBUTIONS TO CO-AUTHORED PUBLICATION: ACTIVITY OF A NOVEL COMBINED ANTIRETROVIRAL THERAPY OF GEMCITABINE AND DECITABINE IN A MOUSE MODEL FOR HIV-1

**I am a co-author on the following publication:**

Clouser CL, **Holtz CM**, Mullet M, Crankshaw DL, Briggs JE, O'Sullivan MG, Patterson SE, Mansky LM. 2012. Activity of a novel combined antiretroviral therapy of gemcitabine and decitabine in a mouse model for HIV-1. *Antimicrobial Agents and Chemotherapy* 56:1942-1948.

My contributions to this publication are as follows:

- Generated data in Figure A2-1
- Contributed text for data generated and edited a complete draft of the manuscript

## Introduction

The efficacies of current anti-HIV-1 therapies are limited by the emergence of drug-resistant virus, which necessitates the development of drugs that exploit novel drug targets. The development of anti-HIV-1 drugs has been relatively rapid and efficient, which has expanded the repertoire of drugs available to HIV-1-infected individuals. The HIV-1 mutation rate represents a potential drug target that has yet to be successfully exploited. Compounds that intentionally increase the mutation rate have been shown to render the virus unable to replicate with enough fidelity to remain infectious, a process termed lethal mutagenesis (36, 59, 85, 132, 209). Although a number of compounds have been shown to lethally mutagenize HIV-1 and one has even made it to clinical trials, none have yet been efficacious and safe enough to be approved for clinical use.

Our previous work identified a combination of clinically approved drugs, decitabine and gemcitabine, that when used together synergistically inhibited HIV-1 replication in cell culture (50). Furthermore, lethal mutagenesis accounted for at least part of the mechanism of action for this drug combination. Besides having a novel mechanism of action, the concentrations of decitabine and gemcitabine necessary to inhibit viral replication were significantly lower than those required to inhibit cell proliferation, suggesting that these drugs may be clinically relevant (50).

In this study, we examined the efficacy and toxicity of decitabine and gemcitabine in the LP-BM5/murine AIDS (MAIDS) model of HIV-1. LP-BM5, the infectious agent of MAIDS, is composed of three different viruses that, upon infection of C57BL/6 mice, cause a syndrome that mimics the immunodeficiency seen with HIV-1 (112, 154). Important similarities between HIV-1/AIDS and MAIDS include (i) a dependence on

$CD4^+$ T cells for disease initiation, (ii) early onset of hypergammaglobulinemia, (iii) loss of B and T cell responses with disease progression, (iv) splenomegaly, and (v) increased susceptibility to opportunistic infections with disease progression. Besides similarities in disease initiation and progression, the MAIDS model has been validated with anti-HIV-1 drugs such as zidovudine and tenofovir, making it a suitable model for screening potential new anti-HIV-1 drugs that have shown promise in a cell culture system (67, 165, 190, 191). We previously used this model to demonstrate the antiretroviral activity of gemcitabine. Specifically, the data demonstrated that both 1 and 2 mg of gemcitabine per kg of body weight decreased disease progression; however, 2 mg/kg was superior in measurements of spleen and lymph node histopathology (49). In preliminary studies, we examined the antiviral activity of decitabine using the MAIDS model and found that 0.15 mg/kg of decitabine decreased disease progression in MAIDS mice, while 0.1 mg/kg had minimal antiviral activity (data not shown). In this study, we hypothesized that using a dose of gemcitabine (1 mg/kg) in combination with a dose of decitabine (0.1 mg/kg) that would be suboptimal if used individually would significantly decrease disease progression using measurements such as provirus levels, spleen weights, and spleen and lymph node histopathology.

Although the MAIDS model is not a perfect model for HIV/AIDS, we show here that the similarities between the potency of decitabine and gemcitabine against HIV and murine leukemia virus (MuLV) in cell culture validate the use of this model for examining the efficacy and toxicity of decitabine and gemcitabine. Here we demonstrate the in vitro and in vivo activities of decitabine and gemcitabine using the MAIDS model. In cell culture, both gemcitabine and decitabine potently inhibit MuLV replication. In

vivo, the combination of decitabine and gemcitabine inhibited progression of MAIDS at doses that did not demonstrate toxicity, as determined by changes in body weight and liver histopathology. Additionally, certain indicators of disease progression indicate that the drug combination was more effective than the individual drugs, as would be predicted on the basis of the cell culture data.

**Results**

*The combination of decitabine and gemcitabine inhibits MuLV infectivity in cell culture.*

Our previous studies showed that decitabine and gemcitabine have potent anti-HIV-1 activity in cell culture (50) and that gemcitabine alone has anti-MuLV activity using the MAIDS model (49). To determine if the LP-BM5/MAIDS model was an acceptable model with which to examine the efficacy and toxicity of decitabine and gemcitabine, we first examined the anti-MuLV activities of the drugs alone and in combination in cell culture. To do this, MuLV expressing GFP was pseudotyped with vesicular stomatitis virus glycoprotein G (VSV-G) and then used to infect target cells that had been pretreated with decitabine, gemcitabine, or the combination of the two. Flow cytometry was then used to determine the percentage of infected cells. Figure II-1A and B show that both decitabine and gemcitabine potently inhibit MuLV infectivity, with $EC_{50}$s in the nanomolar range. Figure A2-1C further demonstrates that the antiviral activity of decitabine and gemcitabine was potentiated when they were used in combination. Although the potentiation was not as great as what is seen with HIV-1 (50), similarities in the potencies of these drugs (Table A2-1) suggested that the MAIDS model would be a valid model for examining the efficacy and toxicity of these drugs.

*The combination of decitabine and gemcitabine inhibits progression of MAIDS as detected by spleen weight and histopathology.*

The data shown in Fig. A2-1 demonstrate that decitabine and gemcitabine inhibit MuLV infectivity in cell culture. To determine if decitabine and gemcitabine inhibit MuLV in vivo, mice were infected with LP-BM5 and treated daily with saline (control), a low dose of decitabine, a low dose of gemcitabine, or low doses of both decitabine and gemcitabine (Table A2-2). The dose of gemcitabine was chosen on the basis of our previously published dose-response analysis performed in the MAIDS model (49). The dose of decitabine was chosen on the basis of our preliminary studies showing that 0.15 mg/kg decitabine decreased the ratio of spleen weight to body weight, while 0.1 mg/kg had minimal effect on the spleen weight to body weight in mice with MAIDS (data not shown). These low doses that demonstrated minimal anti-MuLV activity were chosen so that any potentiation of the drugs when used in combination would be easily observed, as we have previously described (50). Infection of mice with LP-BM5 is characterized by lymphoproliferation, which is detected histologically as a marked expansion of white pulp in the spleen. This expansion also leads to splenomegaly, which is another indicator of disease progression (232). Thus, to examine the extent of disease progression, spleens were removed and weighed immediately after sacrifice. The effect of drug treatment on disease progression was examined by comparing the ratio of spleen weight to body weight. As expected, infected but untreated animals demonstrated splenomegaly, as indicated by the high ratio of spleen weight to body weight compared to the uninfected animals (Fig. A2-2). Treatment with the individual drugs had no effect on the ratio of

spleen weight to body weight, whereas the combination of decitabine and gemcitabine significantly decreased spleen weight compared to the infected but untreated group. Spleens from these animals were further examined for histopathological changes associated with LP-BM5 infection. Lymphoproliferation was scored using a previously described system (49, 97). Briefly, LP-BM5 causes proliferation of lymphoid cells, leading to marked expansion of the spleen's white pulp (Fig. A2-3; compare Fig. A2-3A and B). Therefore, the histopathological changes in the spleen are based on the ratio of white pulp to red pulp, which increases with disease progression. Animals with normal spleen histopathology were rated with an N, while animals with increasing lymphoproliferation were given a score of 1, 2, or 3, indicating mild, moderate, and severe lymphoproliferation, respectively. Table A2-3 demonstrates the number and percentage of animals identified as having significant lymphoproliferation (as indicated by a score of 2 or 3). None of the uninfected animals had significant lymphoproliferation (Table A2-3 and Fig. A2-3A), whereas all spleens from the infected but untreated animals had significant lymphoproliferation (Table A2-3 and Fig. A2-3B). Treatment with a low dose of decitabine or gemcitabine alone had little effect on spleen histopathology compared to that for the untreated but infected animals. In contrast, the combination of decitabine and gemcitabine reduced the number of animals identified as having significant lymphoproliferation within the spleen (Table A2-3 and Fig. A2-3C).

*Treatment with the combination of decitabine and gemcitabine decreases MAIDS-associated lesions in the lymph nodes.*

Infection with LP-BM5 markedly alters lymph node architecture, making lymph

node histopathology a useful indicator of disease progression (232). When detected, the lymph nodes were also evaluated for lymphoproliferation, which was scored on the basis of the extent to which the lymph node architecture and content differed from that expected of a normal, uninfected animal (49). Similar to the spleen scoring system, lymph nodes from animals with increasing lymphoproliferation were given scores of 1, 2, or 3, indicating mild, moderate, and severe lymphoproliferation, respectively. As expected, lymph nodes were difficult to detect in the uninfected animals, whereas lymph nodes were enlarged and easily detected in infected but untreated animals. In fact, all of the infected but untreated animals demonstrated moderate to severe lymphoproliferation (Table A2-3). Treatment of animals with low doses of either decitabine or gemcitabine alone had little effect on the percentage of animals with scores of 2 or 3 (Table A2-3). In contrast, the combination of decitabine and gemcitabine reduced the percentage of animals with lymphoproliferation scores of 2 or 3 compared to animals treated with the drugs individually (Table A2-3).

*Effects of decitabine and gemcitabine on provirus levels.*

The decrease in LP-BM5-mediated spleen lymphoproliferation and splenomegaly suggests that treatment with decitabine and gemcitabine reduced disease progression, which is expected to correlate with a reduction in provirus levels in animals treated with the combination of decitabine and gemcitabine. To examine provirus levels in the infected mice, genomic DNA was isolated from spleen sections and proviral DNA was assessed by real-time quantitative PCR using the 18S rRNA gene for normalization. Consistent with the histopathology data, neither treatment with decitabine alone nor

treatment with gemcitabine alone reduced provirus levels compared to those in the infected but untreated animals (Fig. A2-4). In contrast, the combination of decitabine and gemcitabine significantly reduced provirus levels compared to those in infected but untreated animals (Fig. A2-4). Although the average provirus level from animals treated with both decitabine and gemcitabine was lower than the average provirus level from gemcitabine-treated mice, the averages were not statistically significantly different between these two treatment groups.

*Toxicity of decitabine and gemcitabine determined by changes in body weight and liver histopathology.*

Loss of body weight is often used as an indicator of drug-associated toxicity in mice. Therefore, all mice were weighed daily, and the change in body weight during the course of the study was compared among the treatment groups. None of the treatment groups showed a significant loss of body weight during the study, indicating that the drugs were well tolerated (data not shown). Similarly, no significant changes in liver histopathology were seen in animals treated with the combination of decitabine and gemcitabine (data not shown), further supporting the suggestion that these drugs were well tolerated in these animals.

**Discussion**

The high mutation rate combined with the rapid replication of HIV-1 drives the emergence of drug resistance under suboptimal treatment conditions. Resistance to one drug often confers drug resistance to other drugs in the same drug class, emphasizing the

need for new drugs that exploit novel drug targets. Our previous studies demonstrated the anti-HIV-1 activity of two FDA-approved drugs, decitabine and gemcitabine (50), that appear to inhibit HIV-1 replication through lethal mutagenesis, a process where the mutation rate prevents the virus from replicating with enough fidelity to remain viable. In this study, we examined the antiretroviral activity of decitabine and gemcitabine in vivo using the LP-BM5 MuLV model (a murine AIDS model). This is an efficient model with which to examine the in vivo efficacy and toxicity of potential anti-HIV-1 compounds and has been validated with approved anti-HIV-1 drugs (67, 165, 190, 191). The disease induced by LP-BM5, the infectious agent of MAIDS, has striking similarities to that induced by HIV-1, including (i) a dependence on $CD4^+$ T cells for disease initiation, (ii) early onset of hypergammaglobulinemia, (iii) loss of B and T cell responses with disease progression, (iv) splenomegaly, and (v) increased susceptibility to opportunistic infections with disease progression. However, there are notable differences in the diseases induced by MuLV and HIV-1, which emphasizes the need to confirm the anti-MuLV activity in cell culture before moving into the animal model.

The results presented here show that decitabine and gemcitabine inhibited MuLV in cell culture at nanomolar concentrations (Fig. A2-1 and Table A2-1), indicating that the LP-BM5/MAIDS model would be a suitable model for examining the efficacies and toxicities of these drugs in vivo. While the potency of the individual drugs was similar when comparing MuLV and HIV-1, the combination of decitabine and gemcitabine did not synergistically inhibit MuLV replication, as determined by the FIC method of evaluating synergy (78, 91). This is in contrast to HIV-1, where the combination of decitabine and gemcitabine synergistically inhibits HIV-1 replication (50). This

difference may indicate that one or both drugs use a different mechanism of action to inhibit MuLV compared to the mechanism used to inhibit HIV-1. Similar to the cell culture data, the antiviral activities of decitabine and gemcitabine were potentiated in vivo according to certain indices of disease progression, including the histopathology of spleen and lymph nodes. While the average spleen weight (Fig. A2-2) was reduced in animals treated with the combination of decitabine and gemcitabine, the provirus levels were not statistically significantly different from those in animals treated with gemcitabine alone. Similarly, the provirus levels were not statistically significantly different when comparing the animals treated with gemcitabine alone to the animals treated with both decitabine and gemcitabine. This lack of potentiation in vivo may be due to the variation within each group as well as the small number of animals per treatment group, or it may indicate a different mechanism of action when comparing the mechanism for these drugs in MuLV infection to those in HIV infection.

Although MuLV and HIV-1 have similar life cycles and mechanisms of replication, they also demonstrate some notable differences that may explain why the combination of decitabine and gemcitabine does not act synergistically to inhibit MuLV replication in vivo. Since each drug exhibited similar (decitabine) or lower (gemcitabine) potencies when comparing activities against MuLV and HIV-1 (Table A2-1), the difference in the ability of these drugs to work in combination suggests a difference in their antiviral mechanism of action when comparing the mechanism against HIV-1 to that against MuLV.

As previously described (50), decitabine and gemcitabine are proposed to inhibit HIV-1 replication by lethal mutagenesis, a process where the mutation rate is increased to

a level that prevents viral replication with enough fidelity for the virus to remain viable. In the proposed model, decitabine is incorporated into HIV-1 DNA during reverse transcription, where it forms noncanonical base pairs, thereby increasing the mutation frequency, which can be detected experimentally as an increase in G-to-C mutations in the provirus. Gemcitabine's anti-HIV-1 activity is attributed to its inhibition of ribonucleotide reductase, which alters deoxynucleoside triphosphate (dNTP) pools, where it potentiates decitabine in one of two ways: (i) gemcitabine reduces dCTP levels, thereby increasing incorporation of decitabine, a cytidine analog, or (ii) gemcitabine increases mutation frequency simply by altering dNTP pools, as would be suggested by previous studies that showed that alterations in dNTP pools alone increased HIV-1 mutation frequency (17, 149).

While gemcitabine's reduction or alteration of dNTP pools is likely to increase the HIV-1 mutation frequency and/or increase incorporation of decitabine, gemcitabine may have a different effect on MuLV replication. For example, the $K_m$ of MuLV reverse transcriptase (RT) for dNTPs is higher than that for HIV RT, indicating that MuLV is less efficient at reverse transcribing viral DNA under conditions of low dNTP levels. Thus, gemcitabine's reduction in dNTP pools may inhibit MuLV reverse transcription since MuLV is unable to efficiently reverse transcribe under lower dNTP concentrations. In contrast, HIV-1 is able to efficiently reverse transcribe viral DNA in the presence of low dNTP levels, supporting the suggestion that gemcitabine's anti-HIV activity is likely due to an increase in mutation frequency either through alterations in dNTP pools or by increases in the incorporation of decitabine. Further biochemical studies will be performed to determine if gemcitabine's anti-MuLV activity is due to inhibition of the

enzymatic activity of reverse transcriptase.

While the antiviral mechanism of gemcitabine is likely to differ when comparing MuLV to HIV-1, we do not believe that the antiviral activity of decitabine would differ between the two viruses. Decitabine is a cytidine analog that is incorporated into viral DNA during reverse transcription. Since the $EC_{50}$s of decitabine for MuLV and HIV-1 are similar (Table II-1) and decitabine does not have an altered sugar moiety (a common feature of nucleoside RT inhibitors that are less potent in MuLV) (29, 210), there is no reason to believe that decitabine would have a different affinity for MuLV RT than HIV-1 RT.

Although both decitabine and gemcitabine are used as anticancer drugs, the concentrations needed to inhibit viral replication are significantly lower (100 to 1,000 times lower) than the concentrations needed to inhibit cell growth in vitro. Our results show that the doses needed to inhibit viral replication in vivo were well tolerated, with no signs of toxicity. The dose of gemcitabine used in cancer chemotherapy in humans is 1,000 mg/$m^2$ once per week for 7 weeks with a 1-week break (193). In contrast, the dose of gemcitabine used in this study, 1 mg/kg, is equivalent to 3 mg/$m^2$ in humans when using the body surface area method to convert mouse dosing to human dosing (183). The standard decitabine dosing is 15 mg/$m^2$ every 8 h for 3 days, with this cycle repeated every 6 weeks for a minimum of 4 cycles. In contrast, the dose of decitabine used in this study, 0.1 mg/kg, is equivalent to 0.3 mg/$m^2$, which, like the dose of gemcitabine, is well below the therapeutic dose used to inhibit cell growth in humans.

One limitation for the use of decitabine and gemcitabine as antiretrovirals is that both are currently administered intravenously. However, Eli Lilly has developed a

gemcitabine prodrug that is currently in clinical trials. Although no decitabine prodrugs are currently in clinical trials, it has been demonstrated that with appropriate oral dosing, it is possible to achieve pharmacologically relevant plasma concentrations for cancer treatment (126). Since the dose of decitabine needed to inhibit HIV replication is expected to be lower, it is possible that a prodrug would not be needed, although without a prodrug, it is likely that decitabine would have to be taken multiple times per day, which is not reasonable for anti-HIV drugs.

In summary, the findings presented here indicate that the combination of decitabine and gemcitabine has potent antiretroviral activity in vivo and ex vivo using the LP-BM5/MAIDS model. These findings, along with previous HIV-1 studies in cell culture with gemcitabine and decitabine, suggest that this drug combination has broad antiretroviral activity that is well tolerated at doses that exhibit antiviral activity.

## Materials and Methods

*Materials.*

U373-MAGI-CXCR4$_{CEM}$ cells (93, 226) and SC-1/MuLV LP-BM5 cells (42, 97, 154) were obtained through the NIH AIDS Research and Reference Reagent Program. 293T cells were obtained from the American Type Culture Collection (ATCC). C57BL/6 mice were purchased from Jackson Laboratories. Dulbecco modified Eagle medium (DMEM) was purchased from MediaGrow. Gemcitabine and decitabine were obtained from Carbosynth (Berkshire, United Kingdom). The IgM enzyme-linked immunosorbent assay kit was from Assay Designs (Ann Arbor, MI). The plasmids pCR-DEF and pCR-18S were a kind gift from Mauro Magnani (University of Urbino, Urbino, Italy) and have

previously been described (39). The plasmid pIRES2-EGFP was obtained from Clontech (Mountain View, CA). The plasmids pMIGR1, pJK3, pL-VSV-G, and CMV-Tat were kind gifts from Vineet Kewal Ramani (NCI-Fredrick). The plasmid pMIGR1 is an MuLV vector containing an internal ribosome entry site (IRES)-green fluorescent protein (GFP) element (173); pJK3 contains Gag Pol driven off the HIV-1 long terminal repeat (LTR) (15). pL-VSV-G is also driven off the HIV-1 LTR (15), and CMV-Tat allows transcription off the HIV-1 LTR (15).

*Cell culture.*

SC-1/MuLV LP-BM5 cells and 293T cells were maintained in DMEM containing 10% fetal clone 3 (FC3) serum (HyClone, Logan, UT) and penicillin-streptomycin at 37°C in 5% $CO_2$. U373-MAGI-$CXCR4_{CEM}$ cells expressing the CD4 receptor and the CXCR4 coreceptor were maintained at 37°C in 5% $CO_2$ in selection medium composed of DMEM with 10% FC3, 1 μg/ml puromycin, 0.1 mg/ml hygromycin, and 0.2 mg/ml neomycin.

*Transfection of 293T cells.*

For production of MuLV, pMIGR1 (13.9 μg), pJK3 (6.9 μg), pL-VSV-G (5.54 μg), and CMV-Tat (1 μg) were transfected into 293T cells using the calcium phosphate method. Forty-eight hours after transfection, cell culture supernatant was collected and frozen at −80°C. Similar conditions were used to produce envelope-deficient HIV-1, except that 293T cells were transfected with the envelope-deficient HIV containing GFP, as previously described (50), and the CXCR4–HIV-1 envelope.

*Drug treatment/flow cytometry.*

U373-MAGI-CXCR4$_{CEM}$ cells (62,000) were plated in 12-well dishes 24 h prior to drug treatment. Twenty-four hours later, gemcitabine, decitabine, or the combination of the two was added to the cultures to achieve the concentrations indicated in Fig. A2-1, when taking the total volume (500 μl of medium and 500 μl of virus) into account. Two hours after drug treatment, 500 μl of virus (either MuLV or HIV) was added to each culture such that 20 to 30% of cells were infected in the absence of drug. Twenty-four hours later, the medium containing drug (or dimethyl sulfoxide) was removed and replaced with fresh medium that did not contain drug. Twenty-four hours after the medium change, U373-MAGI-CXCR4$_{CEM}$ cells were collected and analyzed by flow cytometry to determine the percentage of cells expressing GFP. Infectivity was normalized for each individual experiment by setting the infectivity of the untreated cells to 100 for each experiment and then multiplying the data from the other individual treatments by the number used to convert the no-drug-treated cells to 100. The 50% effective concentrations ($EC_{50}$s) were determined in GraphPad Prism software by plotting the log of drug concentration versus the percent infectivity. The data were fit by nonlinear curve fitting using GraphPad Prism. The $EC_{50}$s and 95% confidence intervals were determined in GraphPad using the log inhibitor-versus-response analysis.

*Mice.*

Female C57BL/6 mice aged 8 to 10 weeks were purchased from Jackson Laboratories (Sacramento, CA) and were housed in standard rodent shoebox caging

without a filter top at 22 ± 1°C with a 12-h light/12-h dark cycle, 60% ± 5% humidity, and 12 air changes/h. Mice were fed lab chow and water ad libitum. The experimental protocol was approved by the U.S. Department of Veterans Affairs Medical Center (Minneapolis, MN) IACUC committee (IACUC protocol number 0803A28341).

*Infection of mice with LP-BM5 MuLV.*

LP-BM5 was produced from confluent SC-1 cells by filtering the cell supernatant through a 0.25-μm-pore-size syringe. The filtered viral supernatant was maintained at −80°C until inoculation was performed. C57BL/6 mice were inoculated with two intraperitoneal injections of 0.25 ml of viral supernatant or DMEM spaced 3 days apart.

*Drug treatment of LP-BM5 MuLV-infected mice.*

C57BL/6 mice were randomly divided into the groups shown in Table A2-2. All mice were treated daily with either gemcitabine, decitabine, both decitabine and gemcitabine, or phosphate-buffered saline (PBS) for 8 weeks beginning at 1 week postinfection. Animals were weighed daily to achieve proper dosing and to detect changes in body mass due to toxicity or infection. Animals did not receive drug treatment 24 h prior to sacrifice.

*Sacrifice of animals.*

Animals were euthanized 8 weeks after drug treatment. Animals were weighed and given 100 mg/kg ketamine plus 10 mg/kg xylazine prior to blood collection from the submandibular vein. After euthanasia, necropsy was performed to assess gross

abnormalities in any organs. Spleens were removed, weighed, and sectioned for histopathological analysis as well as for the quantification of proviral levels. Samples of lymph nodes (when detected) and liver were collected. Tissues for histopathological analysis were fixed in 10% neutral buffered formalin, routinely processed into paraffin, sectioned at a thickness of 4 μm, and stained with hematoxylin and eosin.

*Spleen and lymph node histology.*

Histological examination and scoring of spleen and lymph nodes were performed by an ACVP board-certified veterinary pathologist (M.G.O.) from the University of Minnesota Masonic Cancer Center Comparative Pathology Shared Resource. The severity of lymphoproliferation was scored using a previously published scoring system (49).

*Determination of provirus levels from spleen.*

At 8 weeks posttreatment, defective BM5 (BM5def) DNA content was assayed from genomic DNA isolated from spleen. Total cellular DNA was isolated using a Roche kit as per the manufacturer's instructions. BM5def DNA was quantified by real-time PCR assay as previously described (39). The amount of BM5def DNA in spleen was calculated by interpolation of the experimentally determined plasmid standard curve and was normalized to the amount of 18S rRNA. SYBR green master mix (Applied Biosystems) containing genomic DNA (1 μl) and 25 pM primers was used. The primers used to detect BM5def were 5′-CCTTTATCGACACTTCCCTTTT-3′ and 5′-TGGCAGAGGAGG AAGGTT-3′. The primers used to detect 18S rRNA were 5′-

GTAACCCGTTGAACC CCATT (forward) and 5′-CCATCCAATCGGTAGTAGGG (reverse). Conditions for amplification of BM5def and 18S rRNA included an initial heat activation of the polymerase at 95°C for 13 min, followed by 40 cycles of 95°C for 15 s and 62°C for 1 min. Samples were then heated to 72°C for 5 min and 95°C for 1 min, prior to performing a melt curve analysis from 55°C to 95°C. Data were used for analysis only when the standard curve for each primer set yielded an $R^2$ value of >0.99 with an efficiency of 90 to 110%, single peaks were observed in the melt curves, and template controls gave no detectable amplification. Provirus levels normalized to 18S rRNA levels were determined in triplicate for each set of reactions, and 2 to 3 independent assays were performed with each sample, depending on the amount of tissue available for analysis.

*Determination of drug synergy.*

The fractional inhibitory concentration (FIC) method was used to determine the presence or absence of drug synergy as previously described (78, 91). Briefly, the FIC for gemcitabine was calculated by dividing the concentration of gemcitabine used in the combination treatment by the concentration of gemcitabine required to achieve the same degree of inhibition when used alone. The FIC for decitabine was calculated using the same equation with decitabine concentrations in place of gemcitabine concentrations. To determine synergy, the FIC of each drug was added to obtain a FIC index. A FIC index of ≤0.7 is indicative of synergy.

*Statistical analysis.*

Comparisons among treatment groups for Fig. 1C, 2, and 4 were analyzed by one-

way analysis of variance (ANOVA), followed by the Tukey-Kramer posttest. Since the treatment groups in this study (see Table A2-2) did not contain the same number of animals, the Tukey-Kramer posttest was used since this posttest allows comparisons among groups containing different numbers of animals.

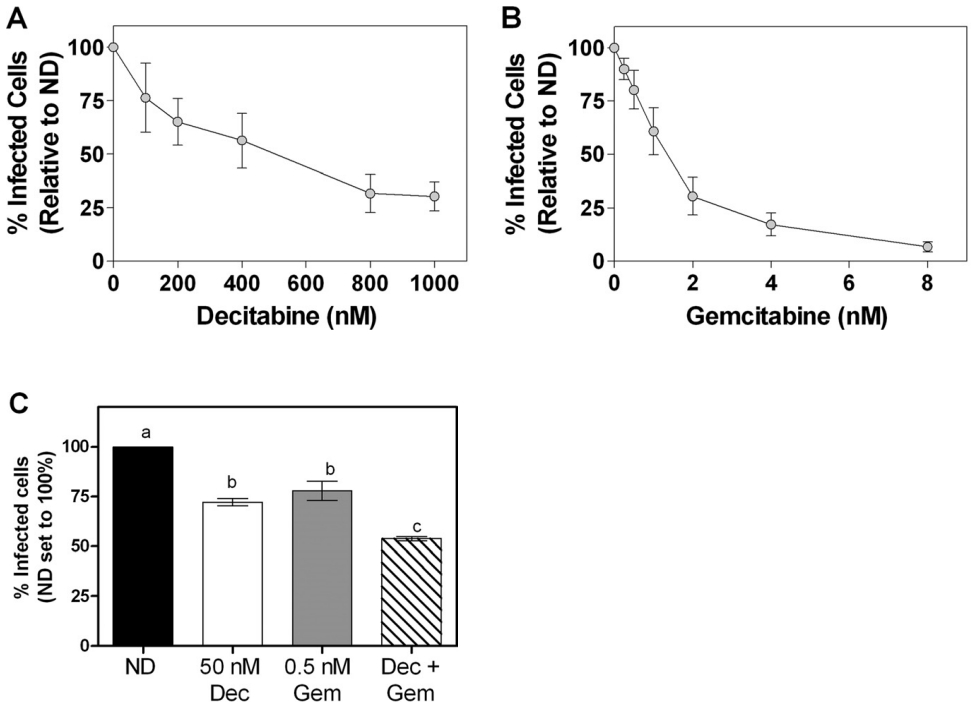



**Figure A2-1. 1`Decitabine and gemcitabine have anti-MuLV activity in cell culture.** MuLV containing GFP was produced from 293T cells and used to infect U373- MAGI-CXCR4CEM cells that were treated with decitabine (A) or gemcitabine (B) at the indicated concentrations. The data represent the average SD of three independent experiments. (C) Potentiation of anti-MuLV activity when decitabine (Dec) and gemcitabine (Gem) are used in combination. U373-MAGI- CXCR4CEM cells were treated with the indicated concentrations of decitabine, gemcitabine, or the combination of the two for 2 h prior to infection. The data represent the average SD of three independent experiments. Statistical significance was assessed using one-way ANOVA with the Tukey-Kramer posttest. Data with the same letters indicate no statistical difference, while data with different letters indicate a statistical difference ($P<0.05$). For

example, the results for the decitabine and gemcitabine treatment groups are not statistically different, while the result for the gemcitabine treatment group is significantly different from that for cells treated with both decitabine and gemcitabine. ND, no drug.

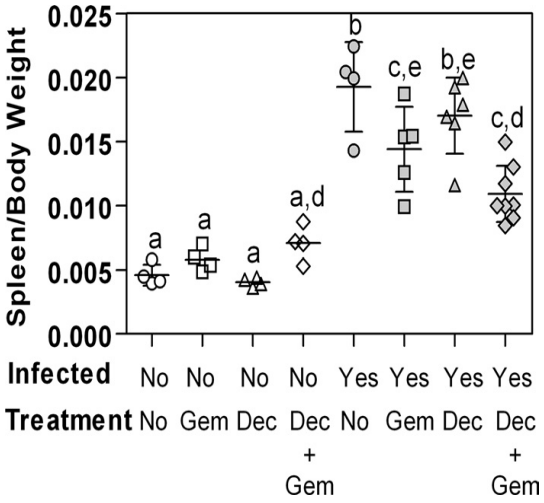



**Figure A2-2. Ratio of spleen weights to body weights from mice infected with LP-BM5 MuLV.** Each symbol (circles, squares, and triangles) represents one mouse. The average SD within each treatment group is shown. Treatment groups that were statistically significantly different from one another, as determined by one-way ANOVA with the Tukey-Kramer posttest ($P<0.05$), are indicated by different letters. For example, the uninfected, untreated group and the group uninfected but treated with Gem are labeled with "a," indicating that there is no statistically significant difference. Dec, mice treated with 0.1 mg/kg decitabine; Gem, mice treated with 1 mg/kg gemcitabine; Dec Gem, mice treated with 0.1 mg/kg of Dec and 1.0 mg/kg Gem. Each data point represents

a single measurement of spleen weight and body weight from an individual animal.

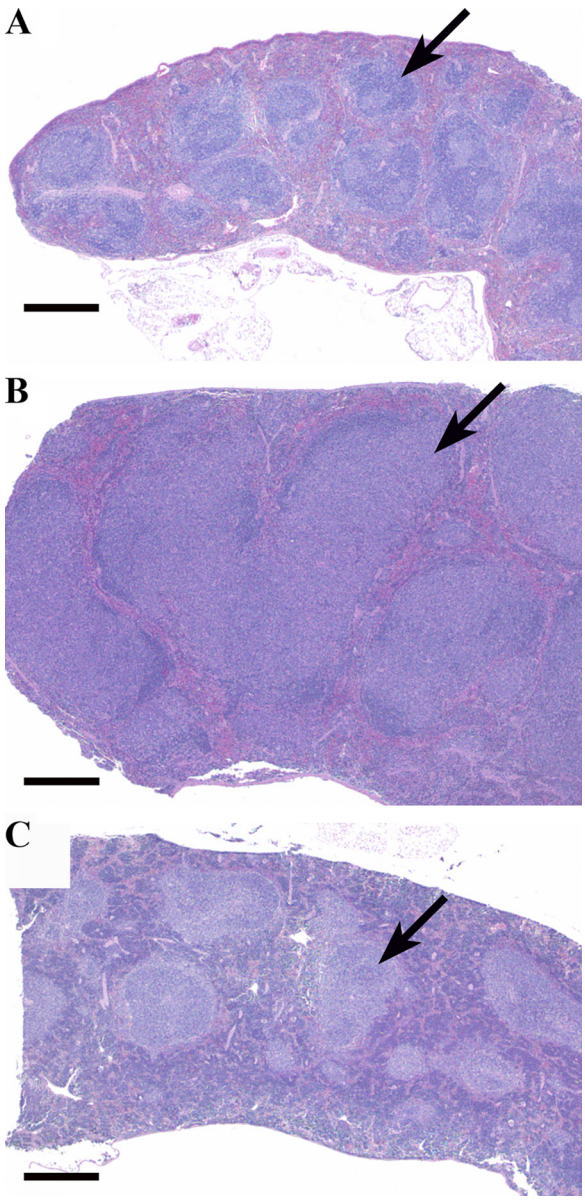



**Figure A2-3. Effect of drug treatment on spleen histopathology.** Spleen tissue was fixed in 10% neutral buffered formalin, routinely processed into paraffin, sectioned at a thickness of 4um, and stained with hematoxylin and eosin. Each panel shows spleen at the same magnification, and arrows indicate the white pulp. (A) Spleen from an

uninfected animal; (B) spleen from an infected but untreated animal showing marked expansion of the white pulp (arrow); (C) spleen from an infected animal treated with 0.1 mg/kg decitabine and 1.0 mg/kg gemcitabine showing marked attenuation of MuLV-induced lymphoproliferation. Bars 500 um.

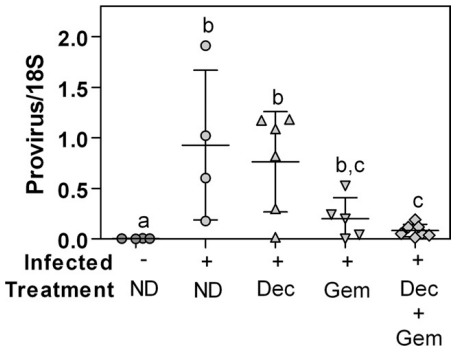



**Figure A2-4. Provirus levels in mice infected with LP-BM5 MuLV.** Genomic DNA was extracted from mouse spleen and quantitative real-time PCR was performed to detect the defective MAIDS provirus. Provirus levels were normalized to the level of the 18S rRNA gene. Each symbol (squares, circles, and triangles) represents one animal. The average SD for each treatment group is shown. The CT (threshold cycle) method was used to assess gene expression. Statistical significance was assessed by one-way ANOVA with the Tukey- Kramer posttest, with a P value of 0.05 designated significant. Treatment groups labeled with different letters are statistically different from one another, while treatment groups labeled with the same letter are not statistically different from one another. For example, the result for provirus from the untreated animals (no drug [ND]) is statistically different from that for provirus from animals treated with both decitabine and gemcitabine, while the result for provirus from animals treated with gemcitabine only is not statistically different from provirus from animals treated with both decitabine and

gemcitabine.

**Table A2-1.** Anti-MuLV and anti-HIV-1 activities of decitabine and gemcitabine in cell culture[a]

| Virus | Drug | $EC_{50}$ (nM) | 95% confidence interval |
|---|---|---|---|
| HIV-1 | Decitabine | 354 | 208–601 |
| MuLV | Decitabine | 361 | 127–1,029 |
| HIV-1 | Gemcitabine | 16.3 | 8.7–31 |
| MuLV | Gemcitabine | 1.6 | 1.1–2.5 |

[a] The data shown are the average from three independent experiments.

**Table A2-2.** Description of treatment groups

| Treatment group | Infection status | Drug treatment (drug dose/day) | No. of animals |
|---|---|---|---|
| 1 | Not infected | No drug | 4 |
| 2 | Not infected | 1.0 mg/kg gemcitabine | 4 |
| 3 | Not infected | 0.1 mg/kg decitabine | 4 |
| 4 | Not infected | 1.0 mg/kg gemcitabine + 0.1 mg/kg decitabine | 4 |
| 5 | Infected | No drug | 4 |
| 6 | Infected | 1.0 mg/kg gemcitabine | 5 |
| 7 | Infected | 0.1 mg/kg decitabine | 6 |
| 8 | Infected | 1.0 mg/kg gemcitabine + 0.1 mg/kg decitabine | 8 |

**Table A2-3.** Effect of treatment on spleen histopathology[a]

| | No. of animals with histopathology score of 2 or 3/total no. of animals (%) | |
|---|---|---|
| Treatment group | Spleen | Lymph node |
| Not infected, no drug | 0/4 (0) | NA |
| Not infected, decitabine + gemcitabine | 0/4 (0) | NA |
| Infected, no drug | 4/4 (100) | 4/4 100 |
| Infected, gemcitabine | 4/5 (80) | 4/5 (80) |
| Infected, decitabine | 6/6 (100) | 6/6 (100) |
| Infected, decitabine + gemcitabine | 2/8 (25) | 2/8 (25) |

[a] See Materials and Methods for scoring system. NA, not enough tissue from these animals was obtained for histopathological analysis.

# APPENDIX III

## DECLARATION OF CONTRIBUTIONS TO CO-AUTHORED PUBLICATION: CONCOMITANT LETHAL MUTAGENESIS OF HUMAN IMMUNODEFICIENCY VIRUS TYPE 1

**I am a co-author on the following publication:**

Dapp, MJ, **Holtz, CM**, Mansky LM. 2012. Concomitant lethal mutagenesis of human immunodeficiency virus type 1. *Journal of Molecular Biology* 419:158-170.

My contributions to this publication are as follows:

- Generated DNA sequencing data in Figure A3-1 and Tables A3-1, A3-2, and A3-3
- Edited a complete draft of the manuscript

## Introduction

The quasispecies model describes how rapidly evolving viruses exist within a mutation–selection balance (35, 77). The model asserts that high viral mutation rates drive the formation and maintenance of the quasispecies and allow for these viruses to readily adapt to a changing environment. However, the quasispecies also presents a significant obstacle in the successful treatment of rapidly evolving viruses such as human immunodeficiency virus type 1 (HIV-1). It has been predicted that a purposeful increase in viral mutation rate would cause an irreversible meltdown of the genetic information (11, 54, 84, 209). This process, termed extinction catastrophe, has been exploited as a novel therapeutic approach, yet findings from short-term cell-culture-based studies fail to predict the shortcomings observed clinically. For example, RNA viruses have been lethally mutagenized in vitro by various nucleoside analogs, including ribavirin, 5-flurouracil, and 5-azacytidine (5-AZC) (59, 206). Specifically, ribavirin was shown to lethally mutagenize poliovirus and hepatitis C virus (53, 54), while 5-flurouracil was shown to be an active viral mutagen against foot-and-mouth disease virus (206). The compound 5-AZC was also demonstrated to lethally mutagenize HIV-1 in cell culture through induction of G-to-C mutations (59). A related compound, KP1212 was shown to lethally mutate HIV-1 in cell culture; however, the compound did not decrease viral loads or increase viral mutation loads in patient samples (94, 158). Similarly, abundant mutations were identified in patient-derived hepatitis C virus, suggesting purposeful mutagenesis by the ribavirin–interferon regimen (56); however, a second study showed only a transient increase in mutation rate in patients on ribavirin monotherapy, indicating that lethal mutagenesis may not be the sole antiviral mechanism (136).

Lethal mutagenesis can be induced not only by drugs but also by the APOBEC3 (A3) family of proteins (21, 46, 95, 127, 140, 144, 196, 218, 236, 238). These proteins have emerged as innate restriction factors that induce targeted hypermutagenesis of viral genomes. While their importance is suggested by the rapid evolutionary expansion of the A3 locus, most APOBEC3 proteins seem to be active against retroviruses and retroelements. However, A3G and A3F exert potent anti-HIV-1 activity through lethal mutagenesis (reviewed in Refs. (5, 188)). Both A3G and A3F, along with A3B, possess the capacity to restrict other retroviral genera, including murine leukemia virus (gammaretrovirus) (21, 95, 140), human T-lymphotrpic virus 1 (deltaretrovirus) (196), foamy viruses (spumavirus) (64), and equine infectious anemia virus (lentivirus) (140). In addition to retroviruses, hepatitis B virus (hepadnavirus) and adeno-associated virus (parvovirus) are also susceptible to members of the A3 family (46, 218).

The mechanism by which A3G hypermutates retroviral genomes has been well established (reviewed in Refs. (47, 96, 139)). Briefly, A3G is packaged into budding virions, after which the virion matures and binds to a target cell. In the target cell, A3G deaminates cytosines (C) present in the single- stranded negative-sense viral DNA during reverse transcription process. The deamination of C leads to uracil (U), and this pre-mutatgenic lesion can template for adenine (A) during plus-strand DNA synthesis rather than guanine (G). The deamination of C by A3G during reverse transcription generates G-to-A mutation signatures in the resulting provirus (95). However, the ability of A3G to mutate the viral genome depends on its ability to overcome viral countermeasures, such as the HIV-1 Vif protein. In a host-specific manner, Vif targets A3 proteins for proteosomal degradation. However, through saturating A3G levels or less stringent Vif

alleles, A3G proteins can gain access to the nascent virions and mutate the viral genome as described above. The ability of A3G to escape Vif is evident in patient samples where signature mutations indicative of A3G have been observed (167, 225). A deaminase-independent mechanism has been proposed for HIV-1, but this model remains controversial (20, 34, 199, 204).

Many of the compounds that lethally mutagenize HIV-1 are C analogs including KP1212, 5-OH-dC, and 5-AZC (59, 94, 132). Competitive replacement by C mutagens could interfere with A3G-mediated deamination. For instance, the kinetics of 5-AZC- and A3G-generated mutations indicate that 5-AZC incorporation into viral DNA precedes the ability of A3G to catalyze cytosine deamination. Replacement of C with 5-AZC may remove potential sites that would otherwise be mutated by A3G. Therefore, compounds targeting C residues may not be the most efficient at inducing lethal mutagenesis in the presence of APOBEC3 proteins. To examine this type of interaction, we investigated mutagen- specific alterations to both mutation spectra as well as the mutational load. Interestingly, our results show that exposure of HIV-1 to both 5-AZC and A3G concomitantly increased the frequency of G- to-A mutations at the expense of G-to-C mutagenesis. Furthermore, the diminution of G-to-C mutations was dependent on A3G catalytic activity. This is the first demonstration for potentiation of the mutagenic effect of a cytosine analog by A3G expression, resulting in concomitant HIV-1 lethal mutagenesis.

**Results**

*Concomitant antiviral effects of 5-AZC and A3G*

A single-cycle vector assay was used to assess the concomitant antiviral effect of 5-

AZC and A3G (Fig. A3-1). This HIV-1 vector system utilizes a dual reporter cassette to monitor both infection efficiency and mutant frequency. Infectious vector virus was produced by transfection of an HIV-1 vector (pHIG) and vesicular stomatitis virus glycoprotein (VSV-G) envelope expression plasmid into previously characterized 293 cell lines that stably express A3G at physiologically relevant levels (192). These cell lines, previously characterized by our laboratory, were used in this study: (1) a 293-A3G null vector control that does not express A3G, (2) a cell line expressing low A3G levels (293-A3G clone 4), and (3) a cell line expressing high A3G levels (293-A3G clone 10) (192). Virus produced from the A3G-expressing cells results in packaging of A3G into the virus particles. These virus particles are used to infect permissive target cells that have been treated with a range of 5-AZC (10–200 μM). Our assay allows for both HIV-1 mutagens to be present during reverse transcription. Any decrease in antiviral activity can be assessed by flow cytometry as a decrease in double-positive murine heat-stable antigen (HSA) + /green fluorescent protein (GFP)+ -expressing cells.

Similar to previous reports (95, 140, 238), A3G decreased HIV-1 replication (Fig. A3-2A), and this decrease was greater as the expression of A3G increased (293-A3G clone 4 versus 293-A3G clone 10) (192). Specifically, vector virus produced from 293-A3G clone 4 reduced viral infectivity by one-third, while virus produced from 293-A3G clone 10 reduced viral infectivity by two-thirds. When permissive target cells were treated with various concentrations of 5- AZC (i.e., 10 μM, 50 μM, or 200 μM), A3G increased 5-AZC's antiviral effect (Fig. A3-2A). For example, virus treated with 50 μM 5-AZC reduced infectivity by 2- fold compared to vector alone; however, this inhibitory effect was potentiated by 3.4- and 5.8- fold with concomitant exposure to 293-A3G clone

4 and 293-A3G clone 10, respectively (Fig. A3-2A). This same trend was observed with concomitant treatment of 200 μM 5-AZC and increasing amounts of A3G, indicating that A3G increased the inhibitory activity of 5-AZC (Fig. A3-2A).

The concept of lethal mutagenesis suggests that decreased infectivity should correlate with an increase in mutation frequency. Therefore, we investigated if the loss of infectivity seen in Fig. A3-2A correlated with an increase in mutant frequency. To do this, we used the single-cycle assay, but the data were analyzed for expression of marker genes. For example, targets cells expressing a single marker gene (i.e., [HSA+/GFP−] or [HSA−/GFP+]) would likely indicate a loss-in-phenotype mutation that occurred during reverse transcription. As indicated in Fig. A3-2B, A3G expression increased viral mutant frequency. Similarly, the mutant frequency was significantly increased when target cells were treated with either 50 μM or 200 μM 5-AZC. Again, the 5-AZC-mediated increase in mutant frequency was further increased by A3G expression. For instance, vector virus treated with 50 μM 5-AZC increased the mutant frequency by 3.2-fold, while concomitant A3G exposure potentiated this effect by 4.3- and 4.6-fold, 293-A3G clone 4 and 293-A3G clone 10, respectively (Fig. A3-2B). Similar trends were found with concomitant A3G exposure in the presence of 200 μM 5-AZC (Fig. A3-2B).

*Concomitant exposure to 5-AZC and A3G alters HIV-1 mutation spectra*

The increase in mutant frequency suggests that viral genomes were heavily mutated by the action of A3G and 5-AZC. Since both mutagens target cytosines and induce distinct mutations, we examined proviruses that had been exposed to either A3G or 5-AZC (EC50 concentration) or to both. Since the resulting proviruses were expected to

have a very high mutational load, primers were designed to minimize PCR bias caused by heavily mutated sequences (as described in Materials and Methods). Table A3-1 shows the cumulative mutation spectra with each mutation type denoted as the percentage of total number of mutations. As expected, A3G levels correlated to an increase in the percentage of G-to-A mutations. As an example, 65% of all mutations identified were G to A when sequencing proviruses exposed to low A3G levels (i.e., 293-A3G clone 4), whereas 70% of mutations were of mutations per viral clone (Table A3-2 and Fig. A3-3A-F). As expected, the G-to-A mutational load increased from 0.03% to 0.35% and 0.37%, when comparing viral clones unexposed to A3G to those exposed to A3G low and A3G high levels, respectively (Table A3-2). The G-to-C mutational load was also increased to 0.31% relative to no drug controls (Table A3-2). Unexpectedly, the G-to-A mutational load was 18% higher ($p = 0.026$) when A3G and 5-AZC exposures were combined, compared to this amount of A3G alone (Table A3-2). Moreover, the concomitant exposure of A3G and 5-AZC also diminished the G-to-C mutational load by 36% ($p=0.013$) compared to drug alone. These results suggest that 5-AZC and A3G may interact to influence the G-to-A and G-to-C mutational load.

Since both A3G and 5-AZC induce mutations by targeting C residues, we next examined viral mutation spectra of combined expression of A3G with 5-AZC treatment. As shown in Table A3-1, viral genomes exposed to the 5-AZC EC50 with increasing levels of A3G significantly decreased the percentage of G-to-C transversion mutations (Fig. A3-3B, E, F). In particular, 61% of the mutations in viruses that were exposed to 5-AZC alone were G-to-C mutations, whereas 41% of the mutations in viruses exposed to both 5-AZC and a low level of A3G were G-to-C mutations. Finally, only 20% of

mutations in viruses exposed to both 5-AZC and a high level of A3G were G-to-C mutations. The general increase in mutational loads suggests that A3G and 5-AZC significantly increased the viral mutation frequency, which is consistent with lethal mutagenesis.

The hotspot prediction algorithm CLUSTERM was used to determine whether 5-AZC-induced G- to-C mutations or A3G-induced G-to-A mutations were influenced by preference sites (i.e., hotspots). As expected by the random substitution of 5-AZC for C bases, the hotspot prediction algorithm found no G-to-C hotspots within GFP. Alternatively, as described in previous A3G studies, the clustering of G-to-A mutations at GG dinucleotides was predicted by CLUSTERM (185, 186). Of the 51 di-G nucleotide positions in the plus-strand DNA of GFP, 26 bases were found to represent A3G preference sites.

*Disparate mutational load with concomitant 5-AZC and A3G exposure*

While mutational spectra analysis offers a population view of the 5-AZC- and A3G-exposed virus, analysis of mutational load can discern differences among individual viral clones. The mutational load was determined for each mutation type of interest (i.e., G to C and G to A) by finding the mean number of mutations per viral clone (Table A3-2 and Fig. A3-3A-F). As expected, the G-to-A mutational load increased from 0.03% to 0.35% and 0.37%, when comparing viral clones unexposed to A3G to those exposed to A3G low and A3G high levels, respectively (Table A3-2). The G-to-C mutational load was also increased to 0.31% relative to no drug controls (Table A3-2). Unexpectedly, the G-to-A mutational load was 18% higher ($p = 0.026$) when A3G and 5-AZC exposures

were combined, compared to this amount of A3G alone (Table A3-2). Moreover, the concomitant exposure of A3G and 5-AZC also diminished the G-to-C mutational load by 36% ($p=0.013$) compared to drug alone. These results suggest that 5-AZC and A3G may interact to influence the G-to-A and G-to-C mutational load.

*A3G catalytic activity is required for reductions in 5-AZC-mediated G-to-C mutations*

Since A3G has been shown to possess catalytic-independent antiviral activity, we next examined the effect of a catalytically inactive A3G on the infection efficiency. It was observed that the A3G_E259Q mutant had no effect on either the efficiency of infection or mutant frequency compared to that of vector only (Fig. A3-4A and B). Importantly, there was no difference in G-to-C mutation spectra under concomitant exposure to 5- AZC and A3G_E259Q compared to that of 5-AZC alone, that is, 57% versus 59%, respectively (Table A3-3). Immunoblot analysis of A3G virion incorporation demonstrated that A3G_E259Q was as efficiently packaged into HIV-1 particles as wild-type A3G, indicating that the lack of efficient incorporation of A3G_E259Q was not responsible for these observations (Fig. A3-4C). Taken together, these data indicate that A3G catalytic activity is necessary in order to observe the decrease in infection efficiency as well as the potentiation of the mutagenic effect of 5-AZC by A3G expression, which results in concomitant HIV-1 lethal mutagenesis.

**Discussion**

Lethal mutagenesis as a therapeutic strategy has gained momentum, especially after ribavirin and KP-1212 gained clinical use (53, 158). Most of the exploratory HIV-1

mutagens target C analogs including the triazole base 5-AZC. Since the APOBEC3 protein subfamily also targets C bases for deamination, these mutagens may exert redundant mechanisms when combined. This would suggest that compounds targeting C bases could be less effective mutagens in the presence of A3G. Alternatively, incorporation of 5-AZC could have secondary effects by altering A3G substrate specificity or target-site architecture. In this study, we examined concomitant exposure of HIV-1 to 5-AZC and A3G. The ability of A3G to induce G-to-A hypermutagenesis provides an example of an evolutionary conserved mechanism that eliminates HIV-1 infectivity by lethal mutagenesis. Editing of retroviral genomes by A3G in the face of concomitant mutagen exposure has not been previously explored.

5-AZC is a first-in-class hypomethylating agent as substitution of N-5 can no longer be methylated (114). The ribonucleoside 5-AZC undergoes anabolic metabolism to the corresponding triphosphate, but a fraction is converted to the nucleotide triphosphate 5-aza-2′-deoxycytidine and is substituted for dCTP in DNA (131, 168). Because aza-pyrimidines are chemically unstable, ring-opened intermediates may have the potential for non-Watson–Crick base pairing, leading to base mispairing (187). In fact, 5-AZC has been implicated in several studies as a mutagen, specifically introducing a rare GC-to-CG transversion type (57, 59, 109, 171). Experimentation with an HIV-1 vector system found that 5-AZC increased the preponderance of only G-to-C mutations, suggesting that the mutagenic product of azacytosine is able to template for C:C mispairings when incorporated during minus-strand DNA synthesis (59).

Intriguingly, the majority of retroviral minus- strand DNA is also transiently single-stranded during replication, and single-stranded DNA (ssDNA) is the preferential

substrate of the APOBEC3 subfamily (95). During HIV-1 reverse transcription, A3G gains access to ssDNA, targeting cytosine residues for deamination to pre-mutagenic lesions (234). These kinetics, of A3G and 5-AZC-based mutagenesis, suggest that access to minus-strand ssDNA by A3G is subsequent to reverse transcriptase incorporation of 5-aza-2′-deoxycytidine during minus-strand DNA synthesis.

The interaction between A3G and the noncanonical 5-azacytosine base could result in two possible obvious outcomes: (1) A3G is unable to engage 5- azacytosine or its ring-opened decomposition products. This would be observed as a decrease in G-to-A mutations with no difference in the number of G-to- C mutations; 5-azacytosine would effectively antagonize A3G deaminase activity. (2) A3G is able to catalyze deamination of 5-azacytosine to 5-azauracil, and this uracil analog results in increased G-to-A mutations at the expense of G-to-C mutation types. However, other mechanisms cannot be excluded at this time, such as the ability of 5- azacytosine (or its ring-opened products) to alter local DNA secondary structure and subsequent ssDNA availability or the influence of 5-azacytosine on A3G processivity or substrate specificity due to close-proximity base interactions, as suggested by Rausch et al (182).

In order to better understand mutation type- specific differences of concomitant mutagen exposure, we sequenced proviral DNAs. The G-to-A mutational load (i.e., number of G-to-A mutations/ proviral clone) was increased up to 18% in virus exposed to both 5-AZC and A3G, as compared to A3G exposure alone. This effect was at the expense of a 36% decrease in the 5-AZC-induced mutational load (i.e., number of G-to-C mutations/sequence). These results suggest a model whereby 5-AZC incorporation into retroviral DNA causes complex interactions with A3G, to inversely shift the G-to-A and

G-to-C mutational loads. Furthermore, the E259Q catalytically inactive A3G, in conjunction with 5-AZC, showed no difference in G-to-C transversions compared to drug alone, suggesting a requirement for fully functional A3G in order to observe the increase in G-to-A mutations at the expense of G-to-C mutations.

One potential concern regarding the sequencing data is biased amplification due to the oligonucleotides selected. Particular attention to primer design helped ensure universal and unbiased PCR. For example, the location of each oligonucleotide was carefully adjusted such that no di-GG nucleotide motif was positioned in the forward primer (and no di-CC nucleotide motif in the reverse primer) in order to eliminate the potential for mis-annealing due to the creation of A3G signatures. Moreover, no more than two G nucleotides were positioned at the forward primer regions (no more than two C nucleotides in the reverse primer region), while these potentially mutable bases were substituted with the degenerate S (50% G and 50% C), to exclude 5-AZC biased amplification. These primers amplified heavily mutated sequences in each treatment group (23 mut/720 bp), indicating that the sequenced DNAs were not biased against any mutation type.

Presently, concomitant use of viral mutagens and chain terminators shows efficacy against RNA viruses (50, 169, 170, 174-176). For example, Perales et al. demonstrated that, in foot-and-mouth disease virus, sequential treatment with a traditional antiviral inhibitor followed by the viral mutagen ribavirin was much more effective in extinguishing picornavirus replication than these compounds used together or either one alone (176). Future progress of mutagen utilization hinges on understanding the mutational constraints within RNA virus population structure. For example, sub-

restrictive editing of retroviral genomes by APOBEC3 proteins has the potential to negatively influence therapeutic intervention, by rapid generation of drug-resistant mutants (116, 157). Furthermore, quasispecies theory predicts that populations will evolve more robust genomes, termed survival of the flattest, in the face of increased mutational load (51, 194). Alternative adaptive strategies include selection of anti-mutator viral polymerases (12, 129, 179) or enhanced discrimination between correct and mutagenic nucleotides (205). Evolution of drug resistance, or increased robustness, of mutagens can threaten such therapeutic approaches; yet, discovery of novel viral mutagens, as well as optimal applications, may lead to alternative therapeutic strategies.

Understanding molecular details of potential mutagens as well as sequence space limitations of specific viral pathogens may provide a rationale for tailored therapeutic intervention rather than lengthy small-molecule screening. Generally, HIV-1 viral mutagens are stealth nucleosides directly utilized by RT during replication to induce site-specific mutations. These compounds are referred to as universal bases because they can mispair with more than one of the canonical Watson–Crick base pairs. For instance, the viral mutagen KP1212 can base pair with either G or A due to a tautormeric shift in the pyrimidine base. 5-AZC, a close derivative to KP1212, induces a unique G-to-C transversion mutational pattern during HIV-1 replication because of its ability to mispair with C bases (59, 109, 187). Even by understanding specific molecular details, viral mutagens still lack a reliable framework to help predict successful treatment outcomes. Many parameters involved in understanding virus population dynamics and genetic diversity are not fully understood, including sequence space, mutational robustness, and effective population size, as well as the natural fitness landscape. Similarly, since

nucleotide base composition among viral genera is known to be quite distinct, purposeful alterations to the mutational bias of a particular virus may pose a greater defect to viral fitness. For example, since HIV-1, like other lentiviruses, has an unusually A-rich genome (19), it is not clear if mutagens that cause more N-to-A mutations (versus A-to-N) are more detrimental to viral fitness.

In summary, we describe concomitant HIV-1 mutagenesis using two unrelated classes of viral mutagens. Our findings indicate a combined, yet intricate, interaction between 5-AZC and A3G. The combined antiviral effect observed indicated that A3G potentiated the mutagenic effect of 5-AZC. Sequencing analysis revealed that the combined mutagenic effect resulted in an increase in the frequency of G-to-A mutations at the expense of G-to- C mutations, suggesting a complex interaction between A3G and 5-AZC upon incorporation into viral DNA. Future studies will provide greater details into the molecular mechanisms involved in concomitant HIV-1 lethal mutagenesis by A3G and 5-AZC.

## Materials and Methods

*Plasmids constructs and cell lines*

The HIV-1 vector pHIG has been previously described (59). Briefly, a ~2.0-kbp dual reporter cassette composed of the HSA, an internal ribosome entry site element, and eGFP was placed in-frame and 3′ to the NL4-3 nef start codon. The G protein of vesicular stomatitis virus (VSV-G) envelope expression plasmid HCMV-G was used to pseudotype virions and was a kind gift from J. Burns (University of California, San Diego). The A3G-expressing cell lines 293-A3G clone 4 and 293-A3G clone 10 were previously

characterized in Sadler et al (192). The APOBEC- expression plasmids pcDNA3.1 + -A3G-3xHA and pcDNA3.1+-A3G_E259Q-3xHA were a kind gift from R. Harris (University of Minnesota). U373-MAGICXCR4 target cells lines were obtained from the AIDS Reagent Program (from M. Emerman).

*Single-cycle transduction and mutant frequency assays*

Vector virus was produced by transient transfection of 293 cell lines with pHIG. Producer cell lines were transfected with 10 μg pHIG and 1 μg VSV-G env plasmids using Genjet ver. II reagent (SignaGen Rockville, MD) as per manufacturer's recommendations. Twenty- four hours post-transfection, viral supernatants were collected and filtered through a 0.2-μm filter. Viral stocks were normalized by transfection efficiency, measured by %GFP-expressing producer cells. Viral stocks were either used to infect target cells or stored at −80°C. The single-cycle infection efficiency for the control A3G null virus was set to 30%. U373-MAGICXCR4 target cells were plated in flat-bottom 12-well dish at a concentration of $6.25 \times 10^6$ cells per milliliter 24h prior to infection. Experiments with the viral mutagen 5-AZC (Sigma- Aldrich, St. Louis, MO) included a 2-h pre-treatment of 1:1000 dilution prior to addition of normalized viral supernatant up to 1 ml total volume. At 24 h postinfection, media were replaced with fresh cell culture media.

*Flow cytometry analysis*

Target cells were prepared for flow cytometry to quantify transduction efficiency (via both GFP and HSA reporter gene expression) and mutant phenotype (via lack of

GFP or HSA reporter gene expression). At 72 h postinfection, cells were collected and stained with 1:250 anti-HSA-PE (BD Pharmingen). Following a 20-min incubation at 4°C, cells were washed 1× in phosphate- buffered saline/2% FC3 and resuspended for flow cytometry. Mutant frequency analysis was determine from the % target cells expressing a single reporter gene relative to the total infected cell population (i.e., %[HSA+/ GFP−] plus %[HSA−/GFP+] divided by % of total infected cells). Mutant frequencies were set relative to virus null for A3G in the absence of drug. Cells were analyzed with a FACScan (BD Biosciences) and CellQuest software. Cells were gated by morphology (forward scatter channel versus side scatter channel) by counting 10,000 cells for fluorescence analysis. Excitation was done at 488 nm; fluorescence channel 1 detected GFP emission at 507 nm, and fluorescence channel 2 detected PE emission at 578 nm. Compensation was set based on single-color controls to eliminate spill-over and re-verified based on the geometric mean of single-color positive to negative-detected populations.

*Proviral DNA sequence analyses*

Approximately 25,000 target cells from each treatment group (+/− A3G and +/− 5-AZC) were collected for total genomic DNA isolation (Roche High Pure PCR Template Preparation Kit) and subjected to nested PCR of the GFP gene in the HIV-1 vector proviral DNA. Primer sets were carefully designed as to remove any bias in amplicon pool considering mutagenic potential of the agents used. Specifically, forward primers did not contain any 5′-GG- 3′ motifs, while reverse primers were absent of 5′-CC-3′ motifs to remove the selection by A3G. Additionally, forward primers were located to

regions containing only two Gs (same for reverse; however, regions contained only two Cs), but these sites were designed to contain a G or C, denoted by S. This approach limited the 5-AZC- induced G-to-C bias that may be excluded from the amplicon pool. Within the pHIG vector, these primer sets would amplify a 944bp region of GFP: outer primer pair, 5′-CTCAATSCCACASCCATA-3′ and 5′-GTSTTSTTTGGGAGTGAA TTAG -3′; inner primer pair, 5′- CTCTCCTCAASCSTATTCAAC-3′ and 5′-GGTATGGSTGATTATGATSTAGAGT-3′. PCR reactions used the Platnium PCR Supermix (Invitrogen), and the initial PCR reaction was purified with GenElute PCR Clean-up Kit (Sigma-Aldrich). Amplicons were verified for correct size and purity by DNA gel imaging prior to ligation into the pGEM-T vector (Promega). Plasmids were transformed into Escherichia coli, and insert-containing-vectors were purified (DirectPrep 96 Miniprep Kit; Qiagen) and sequenced. Sequence alignments were performed using SeqMan of Lasergene 7 (DNAStar, Madison, WI). Only DNA sequences covering the entire GFP gene (720 bp) were used for sequence analysis. This region was chosen because it has been previously used for sequencing analysis of A3G- and 5- AZC-mediated viral mutagenesis (59, 89, 95, 132).

*Immunoblot analysis of transiently transfected A3G-containing particles*

293 cells were transfected with 10 μg pHIG, 1 μg HCMV-G env, and either 5μg pcDNA3.1 + -A3G-3xHA or pcDNA3.1 + -A3G_E259Q-3xHA or 5 μg empty vector, using Genjet ver. II reagent as per manufacturer's recommendations. Twenty-four hours after transfection, cell culture supernatants were collected and passed through a 0.2-μm filter. Next, 1.5 ml of cell culture supernatant was centrifuged at 16,000g for 2 h at 4°C.

Virus particle pellets were resuspended in 2.5 µl of 10× RIPA buffer, and 2.5 µl of a 10× protease inhibitor cocktail was added and incubated on ice for 15 min. Samples were then denatured by the addition of 6× loading dye and boiled for 10 min prior to SDS-PAGE with a 4–20% denaturing gel. Proteins were transferred to nitrocellulose and A3G and A3G_E259Q-containing virus particles were was probed with a primary monoclonal antibody, HA.11 (Covance, Emeryville, CA), followed by a goat anti-mouse horseradish peroxidase conjugate (Invitrogen). To detect HIV-1 capsid (p24) protein, we incubated the nitrocellulose membrane with a rabbit anti-p24 antibody (Advanced Biotechnologies, Columbia, MD) followed by a goat anti- rabbit IgG (H + L)-horseradish peroxidase conjugate (BioRad). Detection of protein location on the nitrocellulose membranes was determined using a BioRad ChemiDoc Imager with Quantity One software (version 4.5.2; BioRad).

*Statistical analyses*

All statistical analyses and graphical representation were done using GraphPad Prism version 5.0 (GraphPad Software, La Jolla, CA). Bars graphs of infection efficiency and mutant frequency data were established with the mean±SD. Statistical analysis of mutational load differences was performed using a chi-squared contingency table, comparing the number of mutated sites to the number of non-mutated sites with each mutation type (i.e., G to A and G to C). A3G hotspot prediction was determined using CLUSTERM†‡. This algorithm predicts a hotspot threshold for the number of mutations at a particular site. The threshold number is established by analyzing the frequency distribution of mutations among a predetermined genomic region (186).

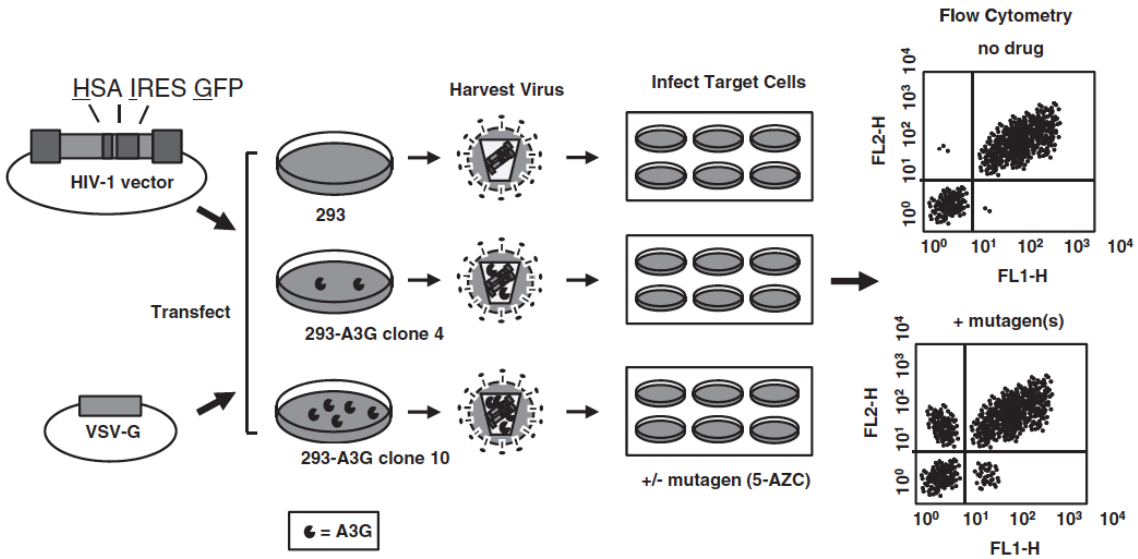



**Figure A3-1. HIV-1 vector assay for analysis of viral mutagens.** The HIV-1 vector, pHIG, was co-transfected with a VSV-G envelope expression plasmid into previously characterized APOBEC3G (A3G)-expressing 293 cell lines (192). Cell culture supernatants were collected and filtered, and virus was used to infect U373-MAGICXCR4 target cells. Following normalization of virus titers, target cells were pre-treated for 2 h with 5-AZC and infected with HIV-1 at a multiplicity of infection of 0.30. Seventy-two hours postinfection, cells were collected for analysis. Abbreviations: HSA, murine heat stable antigen CD24; IRES, internal ribosome entry site; A3G, APOBEC3G; FL1-H, fluorescence channel 1, height of intensity; FL2-H, fluorescence channel 2, height of intensity.

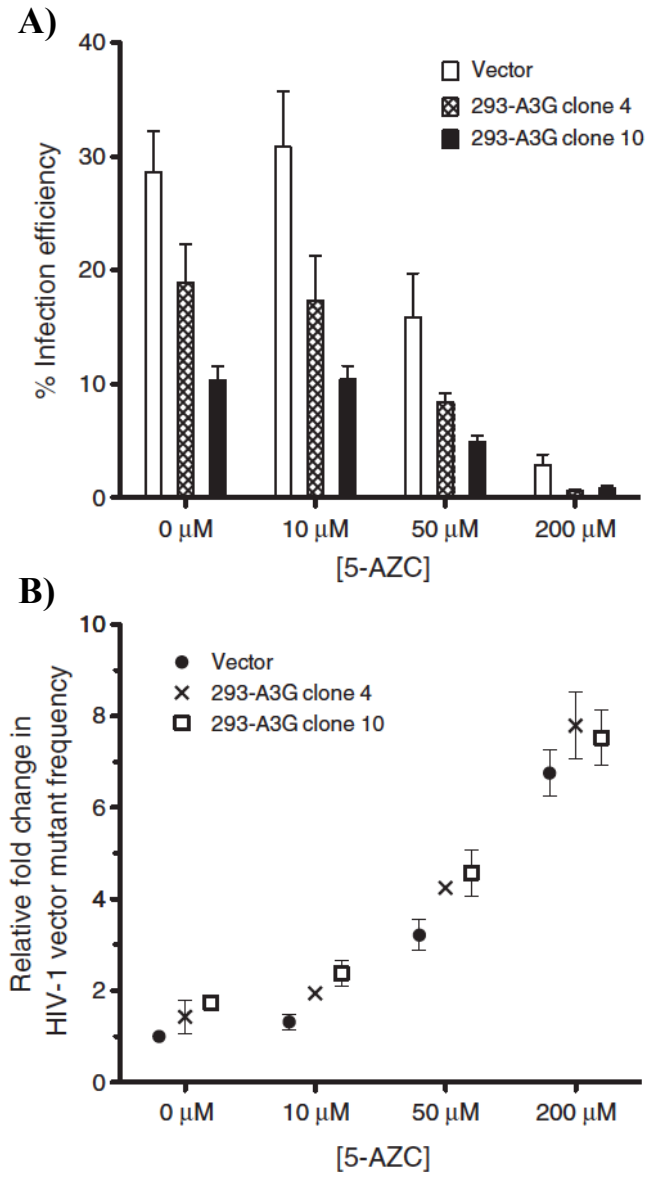
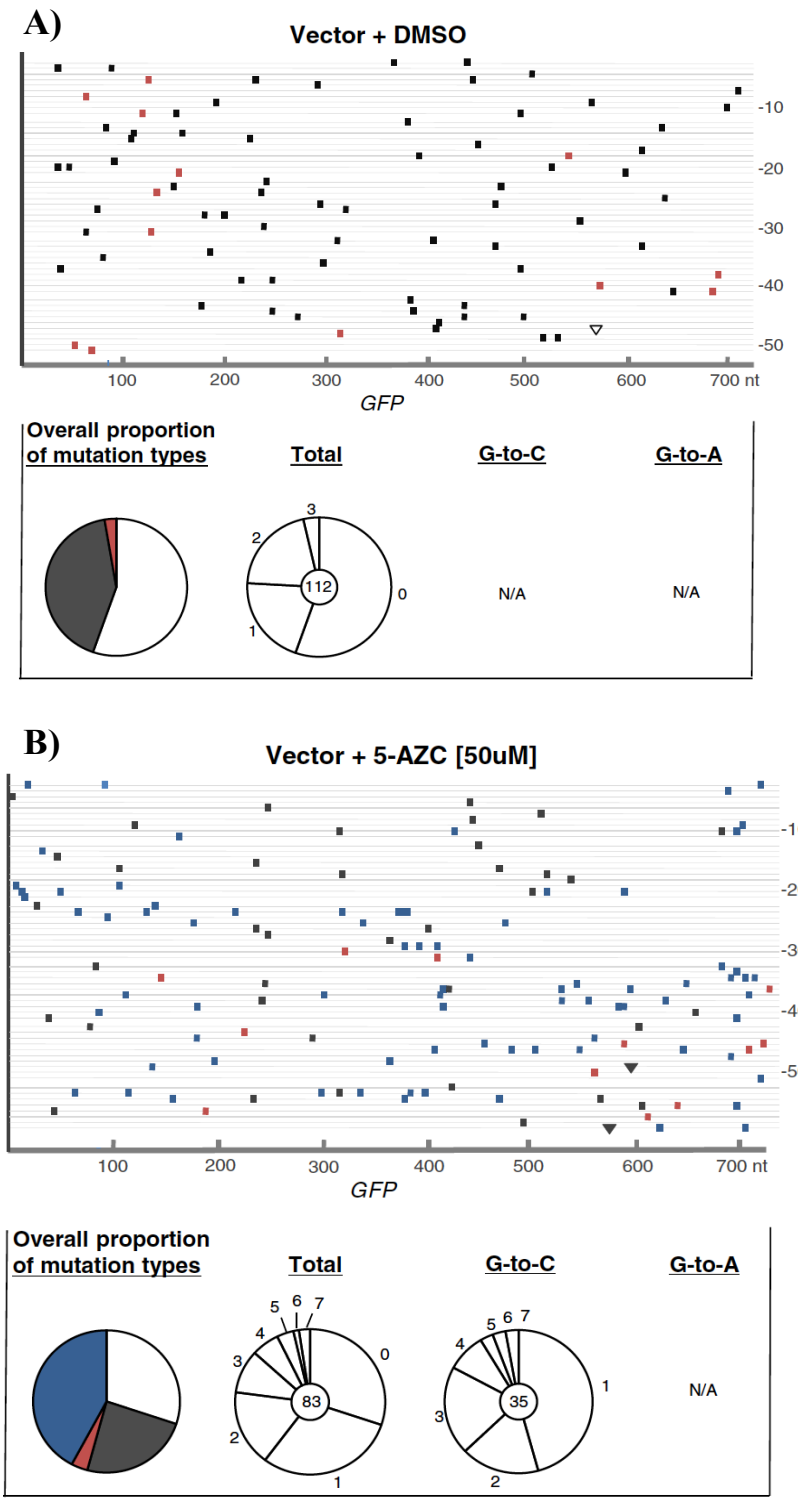
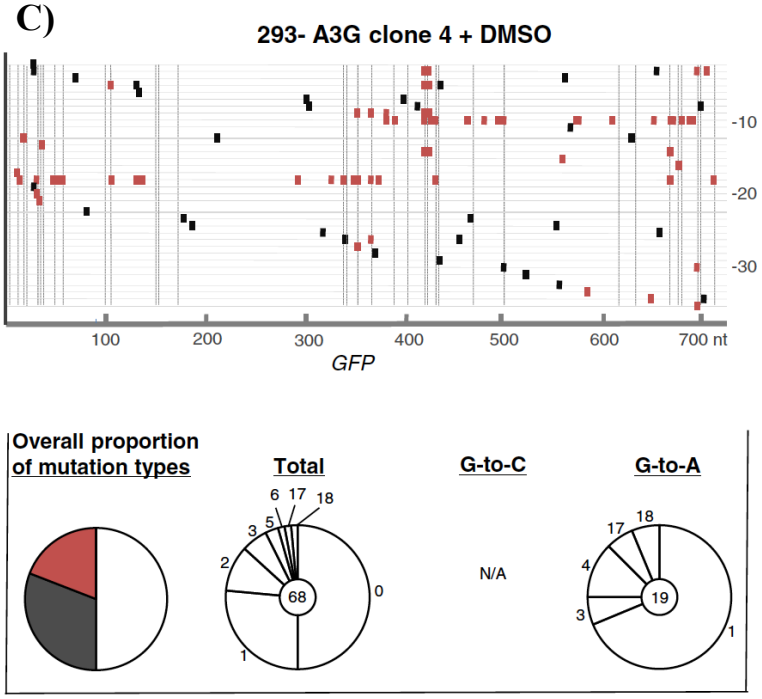



**Figure A3-2. Concentration-dependent inhibition of HIV-1 in the presence of 5-AZC and A3G.** (A) HIV-1 infectivity in the presence of 5-AZC and/or A3G. Virus stocks, produced from the A3G-expressing cell lines 293-A3G clone 4 and 293-A3G clone 10,37 were normalized by transfection efficiency. U373-MAGICXCR4 cells were pretreated for 2 h with the indicated concentrations of 5-AZC prior to infection. Infection levels

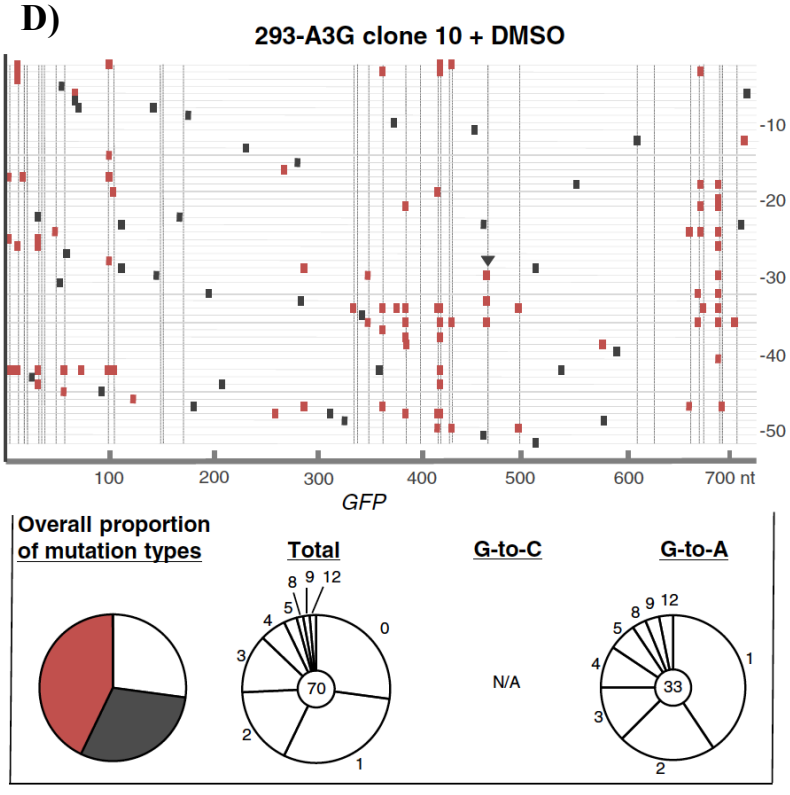
were determined by expression of both the HSA and GFP marker genes. (B) Relative mutant frequency. The mutant frequency is shown relative to the vector + vehicle control. Data represent the mean fold change ± SD from at least three independent replicates. Mutant frequency was calculated as described in Materials and Methods.

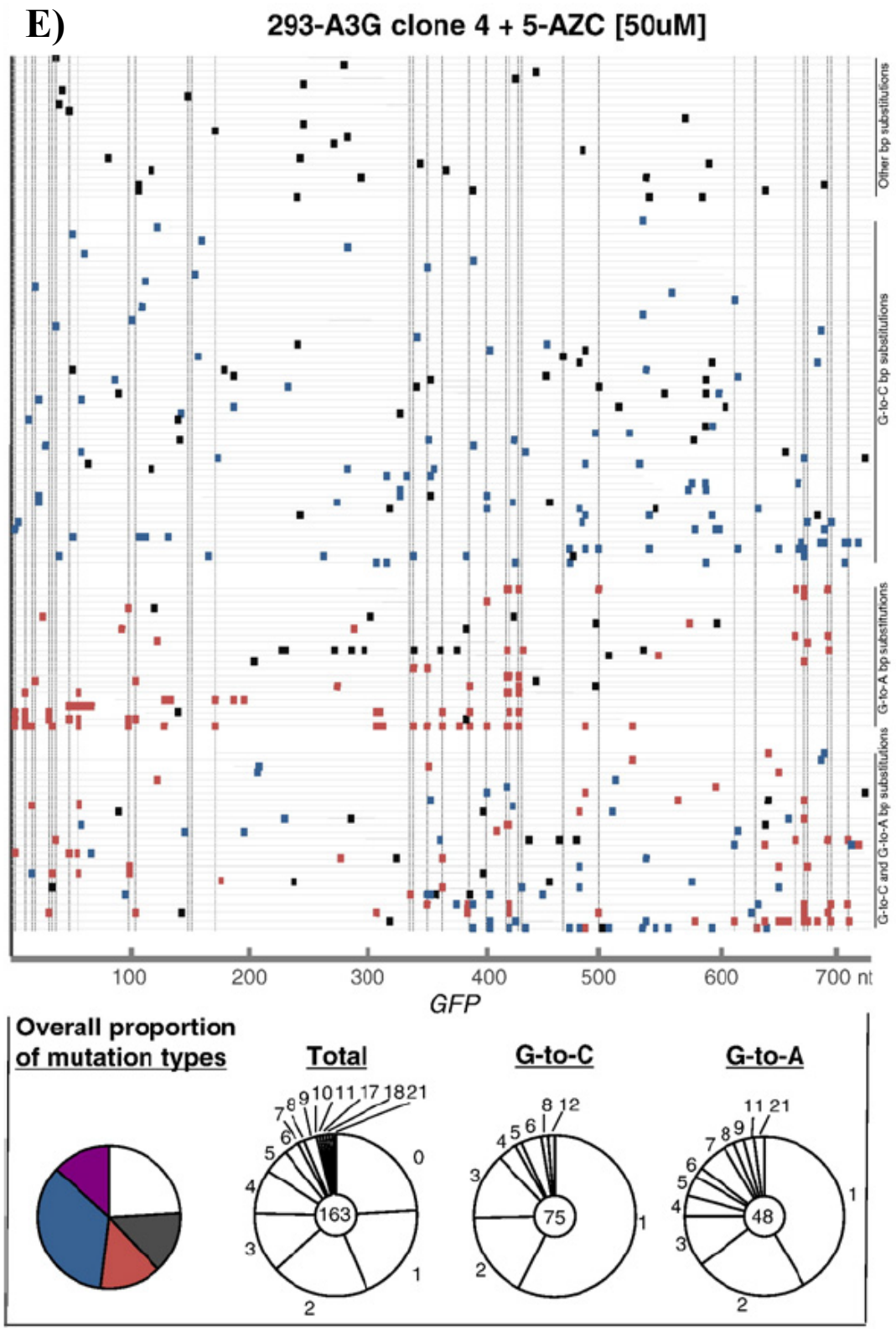

A)
Vector + DMSO
-10
-20
-30
-40
-50
100
200
300
400
500
600
700 nt
GFP
Overall proportion of mutation types
Total
G-to-C
G-to-A
3
2
112
0
1
N/A
N/A
B)
Vector + 5-AZC [50uM]
-1
-2
-3
-4
-5
100
200
300
400
500
600
700 nt
GFP
Overall proportion of mutation types
Total
G-to-C
G-to-A
5
6
7
4
3
83
0
2
1
5
6
7
4
35
1
3
2
N/A

C)


293- A3G clone 4 + DMSO
-10
-20
-30
100
200
300
400
500
600
700 nt
GFP
Overall proportion of mutation types
Total
G-to-C
G-to-A
6 17 18
3 5
2
68
0
1
N/A
17 18
4
19
3
1

D)


293-A3G clone 10 + DMSO
-10
-20
-30
-40
-50
100
200
300
400
500
600
700 nt
GFP
Overall proportion of mutation types
Total
G-to-C
G-to-A
8 9 12
5
4
0
3
70
2
1
N/A
8 9 12
5
4
1
33
3
2


E)
293-A3G clone 4 + 5-AZC [50uM]
Other bp substitutions
G-to-C bp substitutions
G-to-A bp substitutions
G-to-C and G-to-A bp substitutions
100
200
300
400
500
600
700 nt
GFP
Overall proportion of mutation types
Total
G-to-C
G-to-A
163
75
48

**F)**

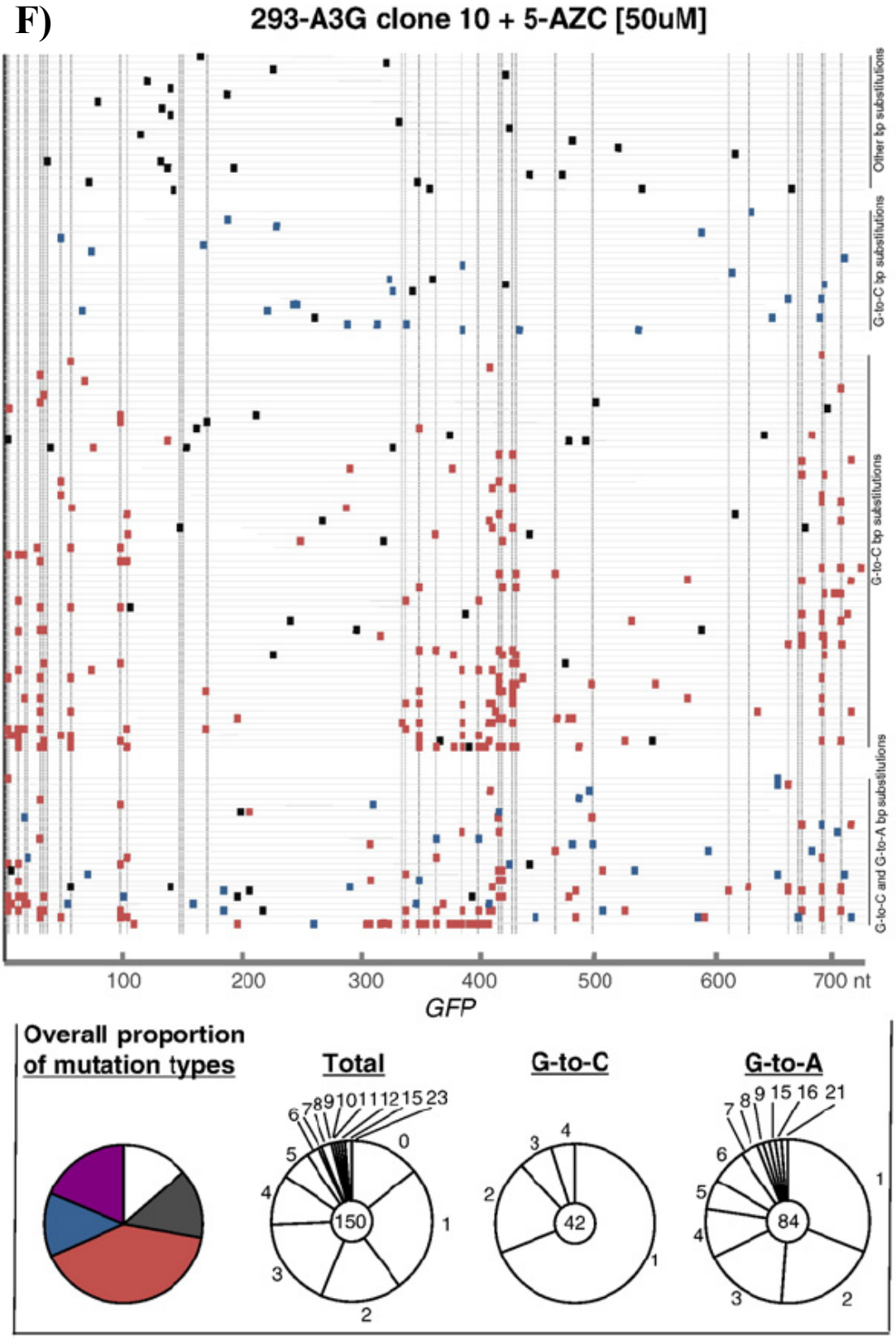
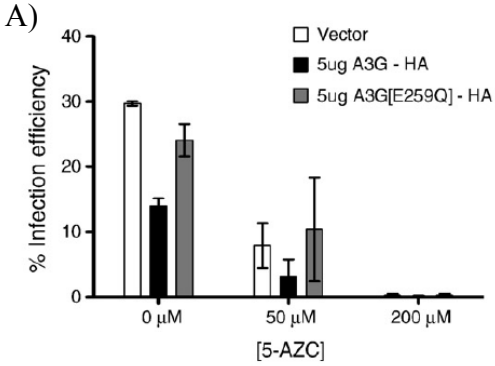
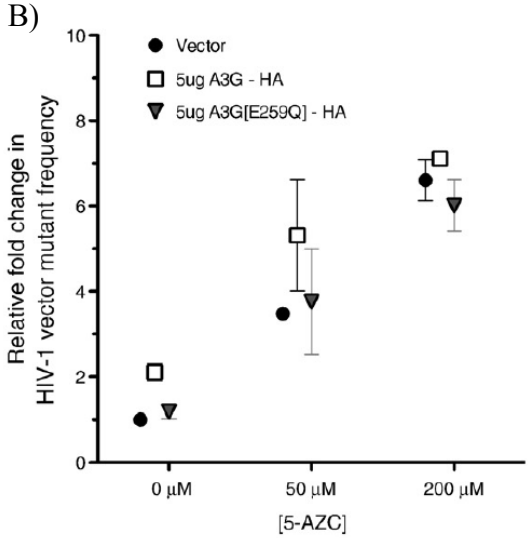

293-A3G clone 10 + 5-AZC [50uM]
Other bp substitutions
G-to-C bp substitutions
G-to-C bp substitutions
G-to-C and G-to-A bp substitutions
100
200
300
400
500
600
700 nt
GFP
Overall proportion of mutation types
Total
G-to-C
G-to-A
150
42
84

**Figure A3-3. Mutation spectra analysis from recovered proviruses.** Total genomic DNA from infected cells was purified, and proviral HIV-1 DNA was amplified, cloned, and analyzed by DNA sequencing. (A) Vector + dimethyl sulfoxide (DMSO), (B) vector + 5-AZC, (C) 293-A3G clone 4 + DMSO, (D) 293-A3G clone 10 + DMSO, (E) 293-A3G clone 4 + 5-AZC, and (F) 293- A3G clone 10 + 5-AZC. Each gray horizontal line represents an individual GFP gene sequence (720 bp) from a recovered provirus (5′ to 3′). Mutation locations are indicated by colored rectangular boxes (G-to-A mutations, orange; G-to-C mutations, blue; other base substitution types, black) or by triangles (insertion mutations, black; deletions, white). Broken vertical lines indicate base locations in the GFP gene sequences that are predicted to be A3G hotspots, as determined using CLUSTERM (as described in Materials and Methods). Below each experimental set (A-F) of GFP gene sequences are summary pie charts that indicate total numbers of G to A, G to C, and other mutation types characterized in experiments. The first column shows colored pie charts and the proportion of recovered clones with no mutations, white; without either G-to-A or G-to-C mutations, gray; with G-to-A mutations, red; with G-to-C mutations, blue; with G-to-A and G-to-C mutations, purple. The other three columns show pie charts indicating the number of mutations per proviral sequence (i.e., 1, 2, 3, etc.) from the total recovered clones (second column), G-to-C-only mutated clones (third column), or G-to-A only mutated clones (fourth column). The total number of proviral clones is indicated by the number at the center of the pie chart. N/A, not applicable.

A)


Vector
5ug A3G - HA
5ug A3G[E259Q] - HA
% Infection efficiency
40
30
20
10
0
0 μM
50 μM
200 μM
[5-AZC]

B)


Vector
5ug A3G - HA
5ug A3G[E259Q] - HA
Relative fold change in HIV-1 vector mutant frequency
10
8
6
4
2
0
0 μM
50 μM
200 μM
[5-AZC]

C)

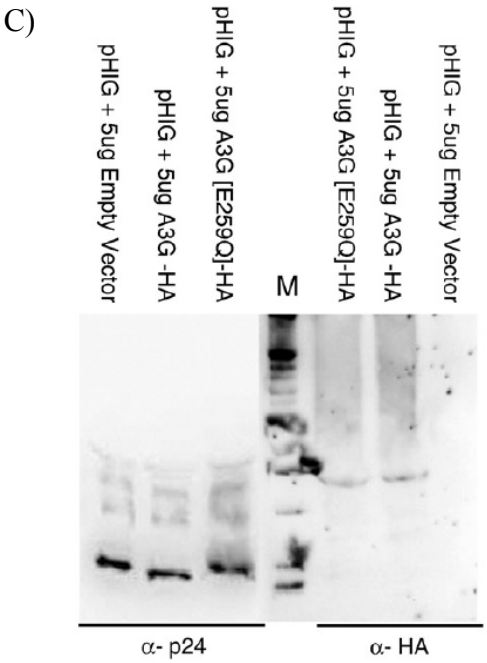

pHIG + 5ug Empty Vector
pHIG + 5ug A3G -HA
pHIG + 5ug A3G [E259Q]-HA
M
pHIG + 5ug A3G [E259Q]-HA
pHIG + 5ug A3G -HA
pHIG + 5ug Empty Vector
α- p24
α- HA

**Figure A3-4. Requirement of APOBEC3G catalytic activity for reducing the frequency of 5-AZC-mediated G-to-C mutations.** 293T cells were transfected with the HIV-1 vector (pHIG), VSV-G envelope expression plasmid, and a A3G- expression vector, the A3G(E259Q) catalytically inactive mutant expression vector, or vector only. Twenty-four hours post-transfection, cell culture supernatants were harvested and used to infect 5-AZC-treated target cells to assay for either viral infectivity (A) or viral mutant frequency (B) determined or used for immunoblot analysis of HIV-1 Gag or A3G (C).

**Table A3-1.** Mutation spectra in the GFP gene of vector proviral sequences

| Mutation from | Mutation to % (±SD) | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Vector+vehicle control[a] | | | | Vector+ 5-AZC [50 μM][b] | | | | A3G clone 4 +vehicle[c] | | | | A3G clone 10 +vehicle[d] | | | | A3G clone 4+5-AZC [50 μM][e] | | | | A3G clone 10+5-AZC [50 μM][f] | | | |
| | A | T | C | G | A | T | C | G | A | T | C | G | A | T | C | G | A | T | C | G | A | T | C | G |
| A to | — | 0 | 2(2) | 39(5) | — | 0 | 0 | 13(9) | — | 0 | 0 | 16(5) | — | 0 | 1(2) | 14(1) | — | 0 | 1(1) | 11(5) | — | 0 | 0 | 6(3) |
| T to | 2(3) | — | 24(8) | 0 | 0 | — | 8(4) | 2(3) | 0 | — | 10(5) | 0 | 0 | — | 9(4) | 1(1) | 0 | — | 5(1) | 0 | 0 | — | 3(2) | 0 |
| C to | 0 | 13(5) | — | 0 | 0 | 6(3) | — | 1(1) | 0 | 9(8) | — | 0 | 0 | 2(2) | — | 1(2) | 1(1) | 7(4) | — | 1(1) | 0 | 5(1) | — | 0 |
| G to | 19(6) | 0 | 0 | — | 4(4) | 5(2) | 61(12) | — | 65(12) | 0 | 0 | — | 70(5) | 1(2) | 0 | — | 28(9) | 4(2) | 41(16) | — | 64(6) | 1(1) | 20(10) | — |

[a] Total number of sequences, 112; total number of mutations, 81; number of nucleotides sequenced, 80,640; mutation frequency $\times 10^{-3}$, 1.035 (±0.136)
[b] Total number of sequences, 83; total number of mutations, 132; number of nucleotides sequenced, 59,760; mutation frequency $\times 10^{-3}$, 1.953 (±0.459)
[c] Total number of sequences, 68; total number of mutations, 95; number of nucleotides sequenced, 48,960; mutation frequency $\times 10^{-3}$, 1.602 (±0.368)
[d] Total number of sequences, 70; total number of mutations, 127; number of nucleotides sequenced, 50,400; mutation frequency $\times 10^{-3}$, 2.203 (±0.335)
[e] Total number of sequences, 158; total number of mutations, 363; number of nucleotides sequenced, 113,760; mutation frequency $\times 10^{-3}$, 3.255 (±0.383)
[f] Total number of sequences, 155; total number of mutations, 468; number of nucleotides sequenced, 111,600; mutation frequency $\times 10^{-3}$, 4.272 (±0.305)

**Table A3-2.** Mutational load in the GFP gene of vector

| | Vector+ 5-AZC [50 μM] | | A3G clone 4+DMSO | | A3G clone 4+ 5-AZC [50 μM][a] | | A3G clone 10+DMSO | | A3G clone 10+ 5-AZC [50 μM][a] | |
|---|---|---|---|---|---|---|---|---|---|---|
| Number of G-to-A mutations | 12 | 0.03% | 48 | 0.35% | 148 | 0.43% | 87 | 0.37% | 287 | 0.48% |
| Total number of G-to-A mutable positions[b] | 41,760 | | 13,680 | | 34,560 | | 23,760 | | 59,760 | |
| % change in G-to-A mutational load | | | | | | +12% ($p > 0.2$) | | | | +18% ($p = 0.026$) |
| Number of G-to-C mutations | 78 | 0.31% | — | 0.00% | 157 | 0.27% | — | 0.00% | 63 | 0.20% |
| Total number of G-to-C mutable positions[b] | 25,200 | | — | | 58,320 | | — | | 30,960 | |
| % change in G-to-C mutational load | | | | | | −13% ($p > 0.3$) | | | | −36% ($p = 0.013$) |

[a] $p$ values determined by $\chi 2$ analysis.
[b] Number of clones multiplied by the 720 nucleotide positions in the GFP target sequence.

**Table A3-3.** Effect of catalytically inactive A3G expression in virus-producing cells on mutation spectra in the GFP gene of vector proviruses

| Mutation from | Mutation to % (±SD) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Vector+5-AZC (50 μM)[a] | | | | A3G E259Q+ DMSO[b] | | | | A3G E259Q+ 5-AZC (50 μM)[c] | | | |
| | A | T | C | G | A | T | C | G | A | T | C | G |
| A to | — | 0 | 0 | 19 (3) | — | 2 (2) | 2 (1) | 19 (0) | — | 0 | 0 | 14 (3) |
| T to | 0 | — | 3 (2) | 5 (2) | 0 | — | 29 (5) | 2 (2) | 0 | — | 7 (1) | 0 |
| C to | 0 | 3 (2) | — | 0 | 0 | 17 (10) | — | 0 | 0 | 10 (3) | — | 1 (1) |
| G to | 10 (4) | 5 (3) | 59 (5) | — | 26 (3) | 2 (2) | 0 | — | 10 (3) | 2 (2) | 57 (10) | — |

[a] Total number of sequences, 36; total number of mutations, 59; number of nucleotides sequenced, 25,920; mutation frequency $\times 10^{-3}$, 2.276 (±0.346).

[b] Total number of sequences, 53; total number of mutations, 42; number of nucleotides sequenced, 38,160; mutation frequency $\times 10^{-3}$, 1.101 (±0.371).

[c] Total number of sequences, 90; total number of mutations, 103; number of nucleotides sequenced, 64,800; mutation frequency $\times 10^{-3}$, 1.590 (±0.312).

## APPENDIX IV: COPYRIGHT PERMISSIONS

**Authorization of:** Holtz CM and L.M. Mansky. 2013. Variation of HIV-1 Mutation Spectra Among Cell Types. Journal of Virology 87(9): 5296-9.

Title: Variation of HIV-1 Mutation Spectra Among Cell Types
Author: Colleen M. Holtz, Louis M. Mansky
Publication: Journal of Virology
Publisher: American Society for Microbiology
Date: May 2013

**OXFORD UNIVERSITY PRESS LICENSE TERMS AND CONDITIONS**

Nov 30, 2013

This is a License Agreement between Colleen Holtz ("You") and Oxford University Press ("Oxford University Press") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

**All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.**

| | |
|---|---|
| License Number | 3279000317356 |
| License date | Nov 30, 2013 |
| Licensed content publisher | Oxford University Press |
| Licensed content publication | Nucleic Acids Research |
| Licensed content title | APOBEC3G cytosine deamination hotspots are defined by both sequence context and single-stranded DNA secondary structure: |
| Licensed content author | Colleen M. Holtz, Holly A. Sadler, Louis M. Mansky |
| Licensed content date | 8 July 2013 |
| Type of Use | Thesis/Dissertation |
| Institution name | |
| Title of your work | STUDIES ON THE ORIGINS OF HIV-1 MUTATION AND GENETIC DIVERSITYSTUDIES ON THE ORIGINS OF HIV-1 MUTATION AND GENETIC DIVERSITY |
| Publisher of your work | n/a |
| Expected publication date | Dec 2013 |
| Permissions cost | 0.00 USD |
| Value added tax | 0.00 USD |
| Total | 0.00 USD |
| Total | 0.00 USD |
| Terms and Conditions | |

**Authorization of:** Clouser CL, Holtz CM, Mullett M, Crankshaw DL, Briggs JE, Chauhan J, VanHoutan IM, Patterson SE, Mansky LM. 2011. Analysis of the ex vivo and in vivo antiretroviral activity of gemcitabine. PLoS One 6(1):e15840.

**Authorization of:** Clouser CL, Holtz CM, Mullett M, Crankshaw DL, Briggs JE, O'Sullivan MG, Patterson SE, Mansky LM. 2012. Activity of a novel combined antiretroviral therapy of gemcitabine and decitabine in a mouse model for HIV-1. Antimicrob Agents Chemother 56(4): 1942-8.

**ELSEVIER LICENSE**
**TERMS AND CONDITIONS**

Nov 30, 2013

This is a License Agreement between Colleen Holtz ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

**All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.**

| | |
|---|---|
| Supplier | Elsevier Limited<br>The Boulevard,Langford Lane<br>Kidlington,Oxford,OX5 1GB,UK |
| Registered Company Number | 1982084 |
| Customer name | Colleen Holtz |
| Customer address | 18-110 Moos Health Science Tower<br>MINNEAPOLIS, MN 55455 |
| License number | 3278990579737 |
| License date | Nov 30, 2013 |
| Licensed content publisher | Elsevier |
| Licensed content publication | Journal of Molecular Biology |
| Licensed content title | Concomitant Lethal Mutagenesis of Human Immunodeficiency Virus Type 1 |
| Licensed content author | Michael J. Dapp,Colleen M. Holtz,Louis M. Mansky |
| Licensed content date | 8 June 2012 |
| Licensed content volume number | 419 |
| Licensed content issue number | 3–4 |
| Number of pages | 13 |
| Start Page | 158 |
| End Page | 170 |
| Type of Use | reuse in a thesis/dissertation |
| Portion | full article |
| Format | both print and electronic |
| Are you the author of this Elsevier article? | Yes |