Assessment and Analysis of the Restriction of Retroviral Infection by the Murine APOBEC3 Protein

Halil Aydin

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies

of the University of Ottawa in partial fulfillment of the requirements

for the degree of Masters of Science

Department of Biochemistry, Microbiology and Immunology

Faculty of Medicine

Learning the rules is hard work and lengthy but playing the game is exciting and fun.

John Polanyi

Nobel Prize Winner in Chemistry

ABSTRACT

Human APOBEC3 proteins are host-encoded intrinsic restriction factors that can prevent the replication of a broad range of human and animal retroviruses such as HIV, SIV, FIV, MLVs and XMRV. The main pathway of the restriction is believed to occur as a result of the cytidine deaminase activity of these proteins that converts cytidines into uridines in single-stranded DNA retroviral replication intermediates. Uridines in these DNA intermediates disrupt the viral replication cycle and also alter retrovirus infectivity because of the C-to-T transition mutations generated as a result of the deaminase activity on the minus strand DNA. In addition, human APOBEC3 proteins also exhibit a deamination-independent pathway to restrict retroviruses that is not currently well understood. Although the restriction of retroviruses by human APOBEC3 proteins has been intensely studied *in vitro*, our understanding of how the murine APOBEC3 (mA3) protein restricts retroviruses and/or prevents zoonotic infections *in vivo* is very limited. In contrast to humans and primates that have 7 APOBEC3 genes, mice have but a single copy. My study of the function and structure of mA3 revealed that it has an inverted functional organization for cytidine deamination in comparison to the human A3G catalytic sites. I have also found that disruption of the integrity of either of these catalytic sites substantially impedes restriction of HIV and MLV. Interestingly, our data shows that mA3 induces a significant decrease in retroviral activity of HIV and MLVs by exploiting both deamination-dependent and -independent pathways. However, the deaminase activity of mA3 is essential to confer long-term restriction of retroviral infection. My observations suggest that mA3 has dual activities, both deamination-dependent and -independent, that work cooperatively to restrict a broad range of human and animal retroviral pathogens. In the context of the intrinsic immune system, APOBEC3

proteins provide a powerful block to the transmission of retroviral pathogens that very few have found ways to evade.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor Dr. Marc-André Langlois for allowing me to be a part of his lab, who with his wisdom and his continuous support has improved my abilities to become a good scientist. He has consistently provided guidance and insight for both my experiments and my writing which I am extremely grateful.

I would also like to thank my thesis advisory committee, Dr. Ken Dimock and Dr. Andrew Makrigiannis, for their input, advice and suggestions on my project and my thesis.

Finally, I would like to thank Olga Agah, Tara Read, Kasandra Bélanger, Shabnam Rahimi, Maria Rosales Gerpe, and Matt Savoie for their technical assistance and advice during my studies at the Langlois laboratory.

This work was funded by a CIHR operating grant to Dr. Marc-André Langlois.

LIST OF CONTRIBUTIONS

Kasandra Bélanger: generated the catalytically inactive hA3G-E69A/E259A expression vector for infectivity assays and time course analysis in Figure-26, -27, -28, -29, -30, -31, -32, -35, -36, -37, -38, -39.

Maria Rosales Gerpe: Schematic representation of hybrid MLV gammaretroviruses in Figure-40.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Acronym	Definition
aa	Amino acid
AID	Activation-induced cytidine deaminase
AKV MLV	AKR virus murine leukemia virus
AKV-N	AKR virus with Fv1 N-tropism
AKV-NB	AKR virus with Fv1 NB-tropism
APOBEC	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like
APOBEC1	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 1
APOBEC2	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 2
APOBEC3	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3
APOBEC3A	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3A
APOBEC3B	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3B
APOBEC3C	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3C
APOBEC3DE	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3DE
APOBEC3F	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3F
APOBEC3G	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G
APOBEC3H	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3H
APOBEC4	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 4
BLV	Bovine leukemia virus
BST-2	Bone marrow stromal antigen 2/Tetherin
CA	Capsid
CD	Catalytic domain
CD	Cluster of differentiation
CD3	Cluster of differentiation 3

CD4	Cluster of differentiation 4
CD28	Cluster of differentiation 28
CMV	Cytomegalovirus
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
Dut	dUTPase
ECL	Enhanced chemiluminescence
eGFP	Enhanced green fluorescence protein
EIAV	Equine infectious anaemia virus
Env	Envelope
ER	Endoplasmic reticulum
ERV	Endogenous retrovirus
F-MLV	Friend murine leukemia virus
FACS	Fluorescence activated cell sorter
FBS	Fetal bovine serum
FIV	Feline immunodeficiency virus
FL-1	Fluorescence 1 channel
Fv1	Friend virus susceptibility 1
Gag	Group specific antigen
Gp	Glycoprotein
HERV	Human endogenous retrovirus
HIV	Human immunodeficiency virus
HIV-1	Human immunodeficiency virus type-1
HMM	High molecular mass
HRP	Horseradish peroxidase

HTLV	Human T cell leukemia virus
IAP	Intracisternal A particle
IL	Interleukin
IL-2	Interleukin 2
IL-7	Interleukin 7
IL-15	Interleukin 15
IN	Integrase
IRES	Internal ribosome entry site
kDA	Kilodalton
LINE	Long interspersed element
LMM	Low molecular mass
LTR	Long terminal repeat
Lv1	Lentivirus susceptibility 1
M5P MLV	M5P packageable genome of murine leukemia virus
MA	Matrix
mA3	Mouse APOBEC3 protein
mA3[Balb]	Mouse APOBEC3 protein from BALB/c mice
mA3[Balb]ΔE5	Mouse APOBEC3 protein without exon 5 from BALB/c mice
mA3[C57]	Mouse APOBEC3 protein from C57BL/6 mice
mA3[C57]ΔE5	Mouse APOBEC3 protein without exon 5 from C57BL/6 mice
mCAT-1	Murine cationic amino acid transporter 1
MDa	Megadalton
MERV	Murine endogenous retrovirus
miRNA	Micro ribonucleic acid
MLV	Murine leukemia virus

MMTV	Mouse mammary tumor virus
MoMLV	Moloney murine leukemia virus
mRNA	Messenger ribonucleic acid
MusD	Endogenous type D murine LTR retrotransposon
NC	Nucleocapsid
Nef	Negative regulatory factor
nm	Nanometer
ng	Nanogram
NP40	Nonidet P-40
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline with tween 20
PCR	Polymerase chain reaction
PERV	Porcine endogenous retrovirus
PFV	Primate foamy virus
Pol	Polymerase
Pro	Protease
R	Repeat region
Rem	Rev-like nuclear export protein
Rev	Regulation of virion
RNP	Ribonucleoprotein
RoRNP	Ro ribonucleoprotein
Rpm	Revolutions per minute
RSV	Rous sarcoma virus
RT	Reverse transcriptase
Sag	Superantigen

SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SFFV	Spleen focus forming virus
SINE	Short interspersed element
SIV	Simian immunodeficiency virus
SIV mac	Simian immunodeficiency virus of macaques
SIV agm	Simian immunodeficiency virus of African green monkey
SU	Surface protein
qPCR	Quantitative polymerase chain reaction
Tat	Trans-activator of transcription
TE	Transposable element
TfR-1	Transferrin *receptor* 1
TM	Transmembrane protein
Trim5α	Tripartite motif-containing protein 5
tRNA	Transfer ribonucleic acid
TU	Transducing unit
U	Unique region
U3	Unique 3’ region
U5	Unique 5’ region
XMLV	Xenotropic murine leukemia virus
XMRV	Xenotropic murine leukemia virus-related virus
XPR1	Xenotropic and polytropic retrovirus *receptor* 1
Vif	Viral infectivity factor
Vpr	Viral protein R
Vpu	Viral protein U

VSV-G	Vesicular stomatitis virus G protein
WB	Western blot
Wt	Wild-type
µg	Microgram
µm	Micrometer
µl	Microliter

1.0 INTRODUCTION

1.1 Retroviruses

Retroviruses are a distinct family of RNA viruses that can cause infections in a variety of vertebrate species. The discovery of two major human pathogens, HTLV and HIV, drew more attention on infectious retroviruses and allowed our understanding of persistent infections to improve drastically over the last few decades. In addition, genomic analysis revealed that a vast portion of the human genome comprises ancient retroviruses called retroelements. Understanding the impact of these elements on genomic stability and the complexity of mammals expanded our knowledge of human evolution. As a result, the infections by exogenous retroviruses and the genomic alterations mediated by endogenous retroelements put retrovirus research in a unique position to contribute both the pathobiology of human infections and the genomic evolution of the human DNA.

1.1.1 Classification of retroviruses

The recent classification of retroviruses subdivided the family members into seven different groups (13). This classification is based on the evolutionary connections among the family members. The avian sarcoma and leukosis viral group, mammalian B-type viral group, murine leukemia-related viral group, Human T-cell leukemia and Bovine leukemia viral group, D-type viral group, lentiviruses, and spumaviruses are the seven groups in the classification and they each contain retroviruses with unique characteristics (13). The first five groups in the classification are also referred to as oncoviruses and they all have simple retrovirus genomic organization (except the Human T-cell leukemia virus (HTLV)-Bovine leukemia virus (BLV) and Mouse mammary tumor virus (MMTV)) (13).

1.1.2 Retroviral Pathogenesis

Retroviruses target a wide variety of animals for their infectious cycle such as sea animals, avian and mammalian species (13). Many other creatures are sensitive to retroviral pathogens, especially primates and mice (13). Retroviruses target different systems and organs during their infectious cycle and cause malignancies, immunodeficiencies, and neurological disorders (13). Moreover, retroviruses may also cause a diverse set of syndromes such as anemia, arthritis, cancer, and osteopetrosis in a wide array of animals (13). Specifically, infection of farm animals may cause a huge economic loss, but most importantly, these pathogens could be transmitted to humans and impose an enormous burden to our health system (13).

Retroviral pathogens have evolved in multicellular organisms, and they tend to use specific cell types for their replication cycle. Retroviruses cause persistent infections and in most cases, these asymptomatic infections translate to only minor effects in the host (13). Infections cause severe damage on multicellular organisms when the virus targets and destroys a single cell type for its replication cycle (13). Retroviruses have developed specific mechanisms to counteract the immune defence mechanisms because persistent infections require long-term interactions with the immune system to maintain the viral life cycle. Animal-to-animal transmission of retroviral pathogens does not usually occur via aerosols and oral pathways, but rather through blood transfusions (fairly limited to humans), sexual contact, and vertical transmission. These modes of transmission are the most important for the virus to persist in its natural host (13). Even though retroviruses usually cause persistent and asymptomatic infections, they may have acute and severe pathologic effects when they

infect different species and become a permanent pathogen in the new host (e.g. HIV/Humans) (13).

Furthermore, the oncogenic features of some retroviruses have also been extensively studied throughout the years. There are many murine models in which retroviruses are used today to understand the molecular mechanisms involved in malignant transformation (13). These studies laid the foundations of molecular tumor biology research (13). Also, oncogene studies, signal transduction assays, and normal cellular growth control analysis are often initially performed with retrovirus-induced malignancies to identify the molecular mechanisms of malignant disorders (13).

1.1.3 Retrovirus genome and morphology

Retroviruses are positive-sense, single-stranded and enveloped RNA viruses. The retroviral RNA is 7-12 kb in size and it is contained in virions 80-100 nm in diameter (13). The retroviral replication cycle is dependent on reverse transcription that transforms the viral genomic RNA into a double-stranded DNA intermediate, which is subsequently integrated into the host's genome as a provirus (Figure-1A) (13, 20). Thus, the genome organization and the morphology of retroviruses have evolved to adapt to the cellular environment in the host organism in order to sustain the viral life cycle.

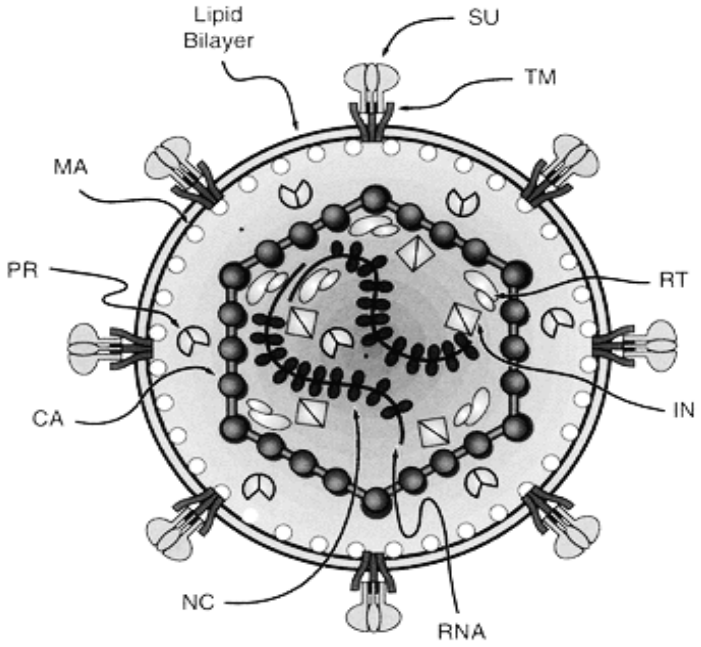
Retroviruses can be distinguished as simple or complex retroviruses in reference to their genomic organization (12, 20, 30). In general, simple retroviruses require only *Gag* (group specific antigen), *Pro* (protease), *Pol* (polymerase) and *Env* (envelope) genes to complete their infectious cycle, whereas complex retroviruses express a variety of additional accessory viral proteins in order to complete their infectious cycle (12, 13, 20). To begin with, the *Gag*

and *Pol* genes are expressed from an unspliced mRNA and the *Env* gene is expressed from a spliced mRNA (20). These are the three essential viral proteins for a sustainable life cycle of both simple and complex retroviruses. The *Gag* gene is responsible for the viral assembly and contains viral core structural proteins; Matrix (MA), Capsid (CA) and Nucloecapsid (NC) whereas the *Pol* gene encodes the viral enzymes; Protease (PRO), Reverse Transcriptase (RT) and Integrase (IN) that are necessary for the viral infectious cycle (20). The *Env* gene encodes the envelope glycoprotein containing two subunits: transmembrane (TM) and surface (SU) proteins (Figure-1A) (4, 20). In contrast, complex retroviruses like the human immunodeficiency virus (HIV) and the simian immunodeficiency virus (SIV) contain accessory genes (*Tat, Rev, Nef, Vif, Vpr, Vpu*) which have distinct roles in the viral infectious cycle that allow for the production and maintenance of new and infectious viral particles (10, 20, 30, 38). Both HIV and SIV belong to lentivirus genus that represents the slow viruses in the Retroviridae family and cause persistent infections that usually results in chronic, progressive disease (13). Two of these accessory genes, *Tat* (trans-activator of transcription) and *Rev* (regulator of virion), encode for regulatory proteins whereas the other four genes, *Nef* (negative regulatory factor), *Vif* (viral infectivity factor), *Vpu* (viral protein U), and *Vpr* (viral protein R), are believed to have evolved in response to the antagonistic effects of host-encoded intrinsic restriction factors (30). The Nef protein enhances the immune evasion by circumventing the adaptive immune system response and increasing the viral spreading, whereas Vif, Vpu, and Vpr proteins counteract intracellular restriction factors by using ubiquitin-dependent proteosomal degradation. Together, these proteins provide a favourable environment for viral replication (Figure-1B) (30). In conclusion, these

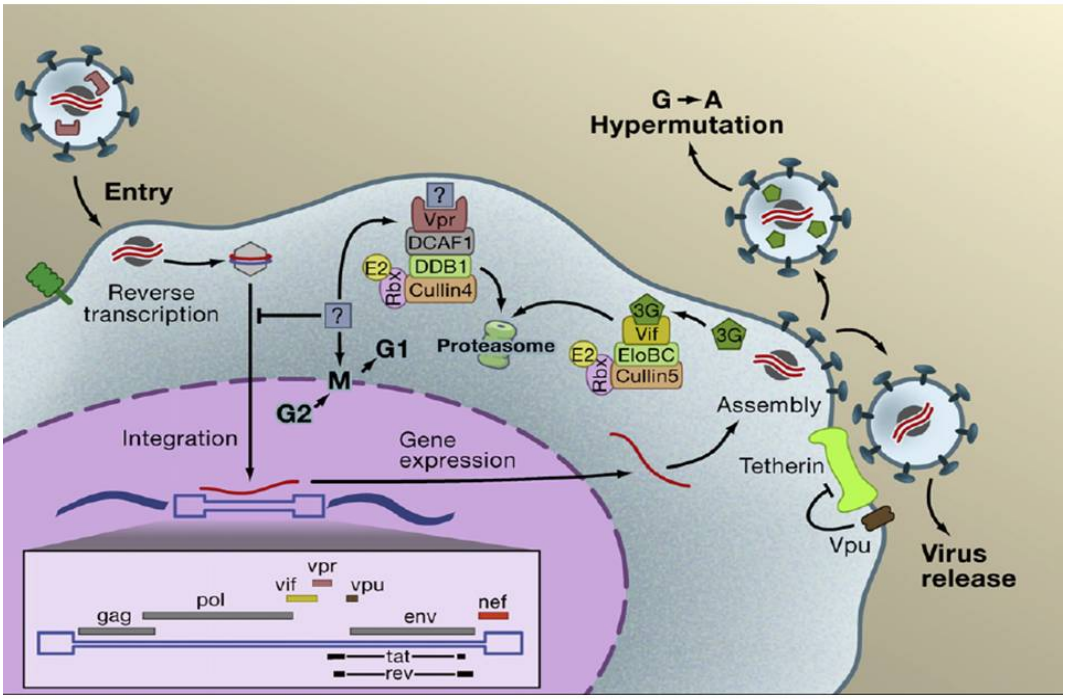
accessory genes are important for facilitating the lentivirus replication cycle and maximizing the infectious potential of complex retroviruses (30).

Figure-1: A) Schematic description of the structure of a retrovirus (13). Transmembrane (TM), surface (SU) subunits are located at the outer surface of the virus whereas capsid (CA), nucloecapsid (NC) and matrix (MA) structural proteins are located in the viral particle and form the inner structure of the particle (13). The viral enzymes, reverse transcriptase (RT), integrase (IN) and protease (PR) are also described in the schematic (13). B) Genomic organization of the HIV-1 lentivirus and the inhibitory role of viral accessory proteins, Vif, Vpu, and Vpr, against the host-encoded intrinsic restriction factors (48). Vif, Vpu, and Vpr proteins counteract intracellular restriction factors by using ubiquitin-dependent proteosomal degradation (48). Together, these proteins provide a favourable environment for viral replication in the target cell and facilitate the release of new viral particles (48).

A)


Lipid
Bilayer
SU
TM
MA
PR
RT
CA
IN
NC
RNA

B)

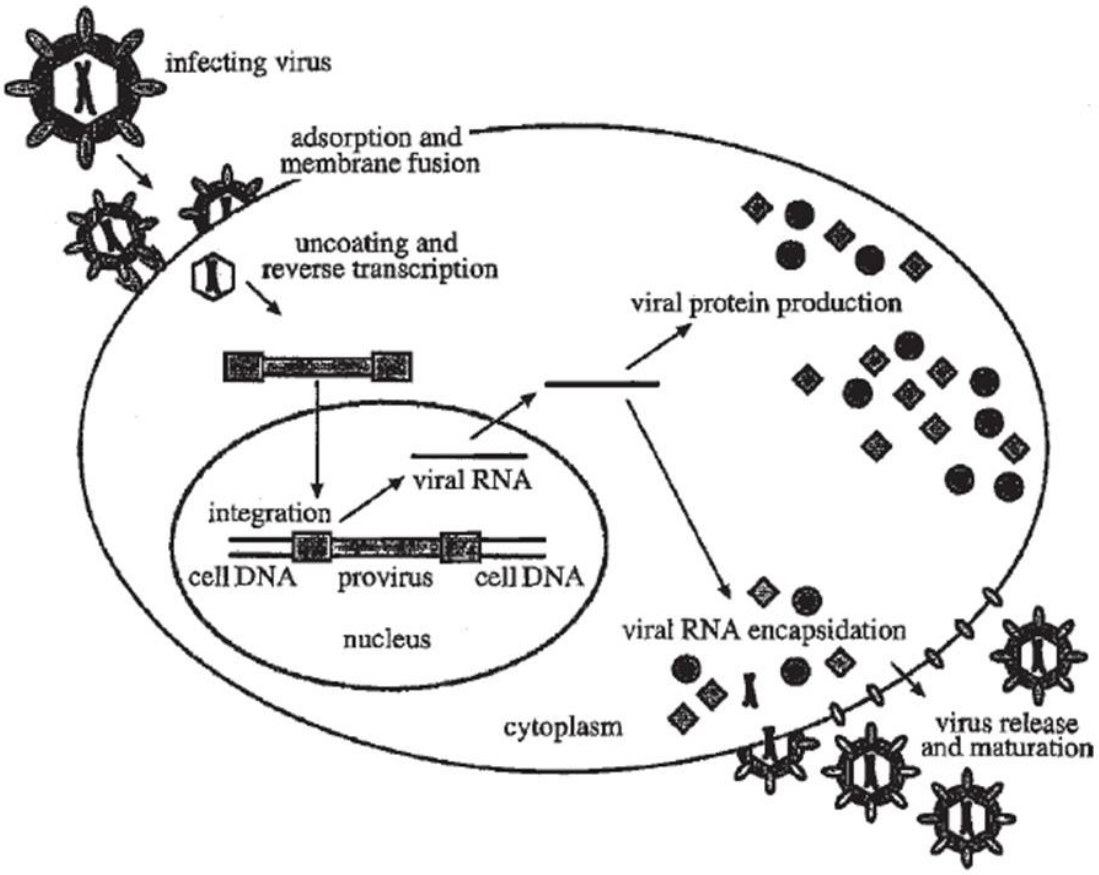

G→A
Hypermutation
Entry
Reverse
transcription
?
Vpr
DCAF1
E2
Rbx
DDB1
Cullin4
?
G1
M
G2
Proteasome
3G
Vif
E2
Rbx
EloBC
Cullin5
3G
Assembly
Integration
Gene
expression
Tetherin
Vpu
Virus
release
vpr
vif
vpu
pol
gag
env
nef
tat
rev

In addition, the 5' and 3' ends of the genomic RNA also have specific segments called long terminal repeats (LTRs) which contain unique (U) and repeat (R) regions (13, 20). LTR segments contain U5, R, and U3 regions at both ends of the viral genome, and these facilitate the replication mechanisms of retroviral pathogens during the reverse transcription (13, 20). The reverse transcriptase produces the double-stranded viral DNA intermediate, which then is translated into the nucleus and integrates into the host cell's genomic DNA, as a provirus (20). The proviral genome has viral promoters and enhancers in the 5' LTR region, which drives the transcription of the viral genome, whereas the transcription termination signal and polyadenylation signal on the 3' LTR region stops the transcription process (20). The U3 region of the provirus contains the transcription control elements, and interestingly, even minor changes in this region may alter the tissue specificity or the replication of the virus, consequently affecting the overall infectious cycle of the virus (13, 20).

1.1.4 Retrovirus replication cycle

Retroviruses have a fascinatingly fast replication cycle compared to other viruses (13, 43). They are able to integrate into the target cell genome within a few hours after viral entry and give rise to new viral particles after 24 hours (43). During the infectious cycle of the virus, the envelope glycoproteins bind to the specific cell receptors required for viral entry after which the fusion of the viral envelope allows for the extrusion of viral core proteins into the cytoplasm. Subsequent uncoating, reverse transcription and integration events lead to the production of new viral proteins and RNAs that eventually move to the cell membrane in order to assemble into new budding particles (Figure-2) (3, 12, 15, 20, 64).

Figure-2: Retrovirus replication cycle in the target cell (20). The envelope glycoproteins bind to the specific cell receptors required for viral entry after which the fusion of the viral envelope allows for the extrusion of viral core proteins into the cytoplasm (13). Subsequent reverse transcription and integration events are essential for the replication of exogenous and endogenous retroviruses, and for the production of new viral particles (3, 12, 13, 15, 20, 64).


infecting virus
adsorption and
membrane fusion
uncoating and
reverse transcription
viral protein production
viral RNA
integration
cell DNA
provirus
cell DNA
nucleus
viral RNA encapsidation
cytoplasm
virus release
and maturation

1.1.5 Gammaretroviruses, their host tropism and distinct features

Gammaretroviruses are classified as simple retroviruses, yet they encode unique features that are essential to their retroviral life cycle. Murine leukemia viruses (MLVs) are such gammaretroviruses that can cause hematopoietic neoplasias in mice (9).

MLV gammaretroviruses can be divided into three different groups according to their host-range tropism (cell type preference) (9). Studies on laboratory mice showed that some murine leukemia viruses can only exploit the mCAT-1 receptor for mouse and rat cell infections, causing leukemia or lymphoma in these hosts (1, 9, 31, 35, 42, 44). The mCAT-1 (murine cationic amino acid transporter 1) receptor is an integral membrane protein responsible for sodium-independent amino acid transportation (44). This group of gammaretroviruses are called ecotropic viruses, since they infect their host species only (9). In their defense, cellular host-encoded restriction factors in mice confer a post-entry block to ecotropic viruses, and thus, restrict their spread to a minimum (1, 9). Alternatively, it has been shown that some gammaretroviruses are not capable of infecting their own host species but they rather infect other species like human, rabbit and cat (9). This group of gammaretroviruses are usually called xenotropic viruses and exploit the XPR1 (Xenotropic and polytropic retrovirus receptor 1) for viral entry (9, 44). The function of the XPR1 receptor is currently unidentified but it is speculated that it could function as a phosphate or amino acid transporter in the cell, similar to receptors used by other gammaretroviruses (9). The last group of gammaretroviruses are identified as polytropic viruses because they are able to infect their own species of origin as well as a variety of other species, therefore, they have the broadest host-range tropism among murine gammaretroviruses with the added ability to exploit the XPR1 receptor for viral entry (1, 9, 31, 44). Importantly, polymorphisms within the XPR1

receptor can define host susceptibility to xenotropic and polytropic murine gammaretroviruses (9).

In addition to the nominal gag protein, which initiates it's translation at an AUG codon, gammaretroviruses also encode a glycosylated form of the gag polyprotein referred to as glycogag (32, 46, 55). Glycogag is an 88 amino acid (aa) glycosylated protein that is translated from an unspliced viral mRNA using an alternative in-frame start codon CUG, which is Leucine, upstream of the AUG gag polyprotein start codon (41). This glycogag segment is also referred to as gPr80, whereas the gag polyprotein, which is the progenitor of viral capsid proteins is called gPr65 (41, 55). The N-terminal region of the glycogag protein contains a signal peptide that is responsible for the glycosylation by directing the protein to the rough endoplasmic reticulum (55). It is then transported to the cell surface where it is cleaved into two proteins (41, 55). The amino-terminal protein fragment changes its configuration to a type II integral transmembrane protein and is reinserted into the plasma membrane (41, 46, 55). Although it is necessary for membrane targeting, glycogag or its cleavage fragments are generally excluded from virion incorporation (41, 46, 55).

For many years, studies on gammaretroviruses focused on glycogag because the role of this protein in the viral infectious cycle was not clear. Even though *in vitro* mutation analysis using Moloney Murine Leukemia Virus (MoMLV) did not show functional defects comparing the mutated glycogag with the wild-type protein, *in vivo* studies indicated that glycogag is necessary for viral replication, viral assembly, and cell membrane budding (41). The conservation of the glycogag segment in the gammaretrovirus genome suggests it is required for sustained replication and release of viral particles in the host, but mutation analysis showed that viruses can replicate without a functional glycogag (gPr80). Studies

removing the glycogag segment have shown this slows the viral kinetics during replication and virion release (41, 46). Moreover, the glycogag of MoMLV directs virion release through lipid rafts, an activity similar to the role of the HIV-1 Nef protein during viral release (55). Thus, one might speculate that glycogag is the equivalent counterpart of the Nef protein within gammaretroviruses (32, 46, 55). APOBEC3 (Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3) (A3) proteins are intrinsic restriction factors that confer the inhibition of retroviral pathogens during the viral replication cycle and interactions in between the glycogag and mA3 protein may shed light on mechanisms involved in the infectious cycle of the murine gammaretroviruses (12, 32). A recent report from Kolokithas *et al.* also examined the glycogag segment of MLV and the inhibitory effects of mouse APOBEC3 (mA3) on glycogag-deleted mutant viruses in mice (32). During the course of their study, mA3 positive and mA3 knockout mice were infected with glycogag-deleted and wild-type MLV virus. Whereas mA3 positive mice only partially restricted the infectivity of wild-type MLV virus, they detected a significant decrease in viral infectivity levels in the absence of glycogag *in vivo,* albeit this study did not investigate the particular mechanisms underlying this antiviral activity (32). Studies have shown that gammaretroviruses interact with the mA3 protein during their replication cycle, however, the potential involvement of the glycogag segment in these interactions remains to be investigated (41, 46). In the context of host innate immunity, the mechanisms behind the glycogag activity are mostly unidentified.

1.1.6 Xenotropic murine leukemia virus-related virus (XMRV)

Xenotropic murine leukemia virus-related virus (XMRV) is a simple gammaretrovirus that was discovered in 2006 (40, 75). It was first thought to be involved in prostate cancer and

chronic fatigue syndrome (CFS), but recent publications have identified the virus as a laboratory contaminate that originated from a distinct recombination event during the generation of the 22Rv1 prostate cancer cell line (40, 60, 75). XMRV shares a high percentage of sequence identity with other endogenous and exogenous murine leukemia viruses and also exploits the XPR1 receptor for viral entry (50, 71, 72). Upon initial identification of XMRV, independent researchers set out to screen CFS and prostate cancer patients but were unable to detect the virus in their samples (21, 39, 71, 72). Studies had also found that intrinsic retroviral restriction factors were effective against the virus (21, 71, 72). After the identification of the true source of the XMRV virus, it is interesting to see how easily new retroviruses can be generated from recombination events between exogenous and endogenous retroviruses (60).

1.1.7 Mouse mammary tumor virus (MMTV)

Mouse mammary tumor virus MMTV is a complex and oncogenic betaretrovirus that was discovered in the 1930s. MMTV is transmitted vertically from the infected mother to the nursing pups through milk and causes mammary epithelial cell tumors (62, 65). MMTV exploits the transferrin 1 receptor (TfR1) and infects activated lymphocytes and dividing mammary epithelial cells *in vivo* (63, 65). The MMTV genome contains various accessory genes to facilitate the viral replication cycle such as *Dut* (dUTPase), *Rem* (Rev-like nuclear export protein), and *Sag* (superantigen). The viral protein Sag is encoded in the 3' LTR and is the most important accessory gene since it controls the infection pathways of the virus in a variety of cell types and tissues (63, 65). When MMTV is acquired through the breast milk, the virus infects the dendritic cells in the gut and subsequently migrates to the lymph nodes where it can then infect lymphocytes (62, 63, 65). When mammary epithelial cells start to

divide from hormonal stimulation during puberty or pregnancy, MMTV can also target these cells for its replication cycle (29, 63). As a result of the initial infection in the mammary gland, the provirus can integrate into the mammary epithelial cell genome and induces the activation of cellular oncogene expression, resulting in mammary tumors (63, 65).

1.1.8 Endogenous retroviruses

Transposable elements (TE) comprise a large portion of the genome in humans (16, 17, 66). TE activity and accumulation correlates with genomic complexity but their uncontrolled movement in the genome can also be harmful for the host (66). Retrotransposons are mobile DNA segments in the genome that exploit reverse transcription and integration mechanisms to replicate through RNA intermediates and then colonize a new region in the genome (16, 66). The classification of retrotransposons subdivides these elements according to the presence or absence of the long terminal repeats (LTR) (16). Retrotransposons with long terminal repeats are called endogenous retroviruses (ERVs) and this group of retrotransposons occupy 8% of the human genome (16, 17, 66). On the other hand, non-LTR retrotransposons account for a whopping one third of the human genome, commonly referred to as long interspersed element 1 (LINE-1), Alu, SVA elements, and other non-LTR retrotransposons (16). The correlation between genomic complexity and retroelement load in the genome may explain the impact of retrotransposon activity on genomic evolution. For instance, transposable elements occupy 3% of yeast, 6% of worm, 15% of flies, 40% of mice and 44% of human genomes, theoretically expanding the evolutionary complexity of the host (66). The enormity of retroelement load is more significant when it is compared with protein coding regions which comprise only 1.5% of the genome (16, 66). In contrast, retroelements may also cause genomic instability, generate insertional mutations or modify the gene

expression in a locus (16). Thus, the presence of retroelements must be balanced to maintain the integrity of the genome, illustrating the important role that intrinsic restriction factors play in counteracting exogenous and endogenous retroviruses and retroelements (66).

The origin of endogenous retroviruses is unclear but it was suggested that they either lost their *Env* gene and endogenized or they acquired an *Env* gene during their transposition and escaped from the cell (66). In addition, endogenous retroviruses are only contained in the vertebrate genome and their classification is organized based on their similarity to the exogenous retroviruses (Table-1) (66).

It is known that some of the retroelements are still active in the human genome and they are responsible for 0.1% of *de novo* mutations (66). Interestingly, retroelements in rodents have a stronger impact on the genome by generating almost 10% of spontaneous mutations. This activity can be attributed to their immensely polymorphic nature (66). The most active murine endogenous retroviruses (ERVs) are MusD/ETn elements and intracisternal A-type particle (IAP) elements. It was shown that intrinsic restriction factors such as APOBEC3 proteins are capable of inhibiting the activity of human LINE-1, Alu, and mouse IAP and MusD elements (66). In summary, endogenous retroviruses may have a strong genetic link to exogenous retroviruses and their interactions with infectious retroviruses may cause specific recombination events that lead to the generation of new retroviruses or alter the genetic structure of old virus (17, 60).

Table-1: Classification of endogenous retroviruses (66). The table was adapted from Rowe *et al.,* 2011 (66).

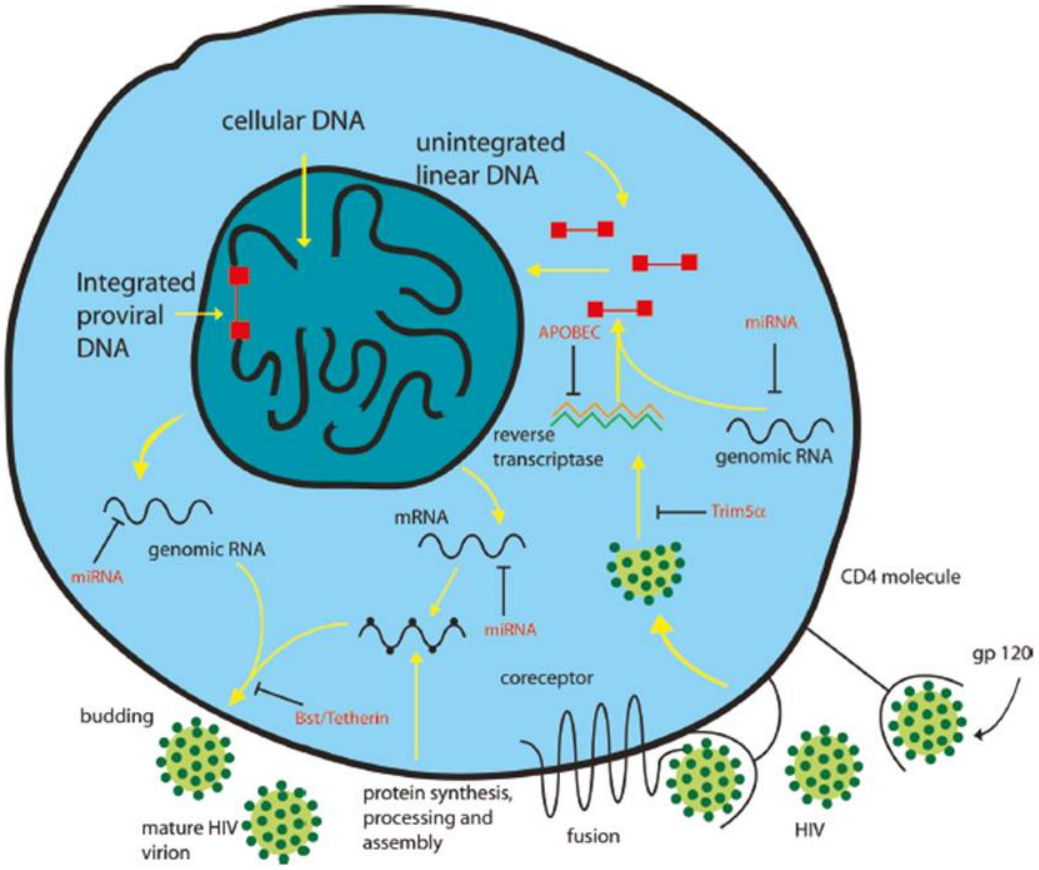
	Class-I	Class-II	Class-III
Mouse ERV	Murine leukemia virus (MLV)	MusD, Intracisternal A particles (IAP), Mouse mammary tumor virus (MMTV)	MERV-L
Human HERV	HERV-H (His), HERV-W (Trp)	HERV-K (Lys)	HERV-L (Leu)
Related XRV genus	Gamma RVs, Epsilon RVs	Alpha RVs Beta RVs Delta RVs Lentiviruses	Spumavirus

1.2 Host-encoded intrinsic restriction factors

All living creatures from bacteria to humans have their own infectious enemies against which they evolve clever defence mechanisms. On the other hand, infectious organisms such as viruses combat these defence mechanisms by hijacking the host metabolic pathways or products to perpetuate their life cycles (64)

In mammals, the innate immune system has evolved cellular restriction factors against retroviruses to impair the rapid retroviral replication cycle and reduce their infectious activity (38, 43). There are at least four major and distinct intrinsic restriction factors effective against retroviral pathogens in mammals (Figure-3) (38). First of all, Fv1 (Friend virus susceptibility 1) (or Lv1 for lentiviruses in primates) is one of the first restriction factors discovered that hinders retroviral pathogens before integration into the target cell genome. Another restriction factor called tetherin (CD317/BST-2) blocks the release of viral budding particles from the cell membrane by tethering newly forming viruses to host cells (1, 25, 30, 38). As well, tripartite motif-containing protein 5 (TRIM5α) is an additional restriction factor that is able to impair retroviral capsid uncoating (27, 43). Lastly, a member of the cytidine deaminase family, APOBEC3, can disrupt the retroviral replication cycle as well as block the activity of retrotransposons (30, 38, 43, 64). These intrinsic restriction factors provide a strong barrier against retroviruses. Their broad activity against old and new strains of retroviral pathogens make them an essential component of the innate immune system, and their expression impairs species-specific as well as inter-species transmission of retroviruses (12, 43, 64).

Figure-3: Host-encoded intrinsic restriction factors and inhibition of retroviral pathogens at different stages of the viral replication cycle (38). Various cellular restriction factors such as tetherin, Trim5α, and APOBEC3 proteins intervene in the retrovirus replication cycle by using their unique restriction mechanisms (38). Tetherin does not allow new forming viral particles to be released from the host cell. APOBEC3-mediated restriction occurs at the reverse transcription microenvironment where A3 proteins deaminate cytidine to uridine on the negative DNA strand and/or affect the processivity of reverse transcriptase. Trim5α interacts with the viral capsid and impairs viral uncoating in the cell (38). In addition, miRNAs have inhibitory effects on viral genomic RNAs and mRNAs during the replication cycle (38).


cellular DNA
unintegrated
linear DNA
Integrated
proviral
DNA
APOBEC
miRNA
reverse
transcriptase
genomic RNA
Trim5α
mRNA
genomic RNA
miRNA
miRNA
CD4 molecule
gp 120
coreceptor
budding
Bst/Tetherin
protein synthesis,
processing and
assembly
fusion
HIV
mature HIV
virion

1.3 APOBEC3 proteins and antiretroviral activity

1.3.1 The discovery of APOBEC3 proteins

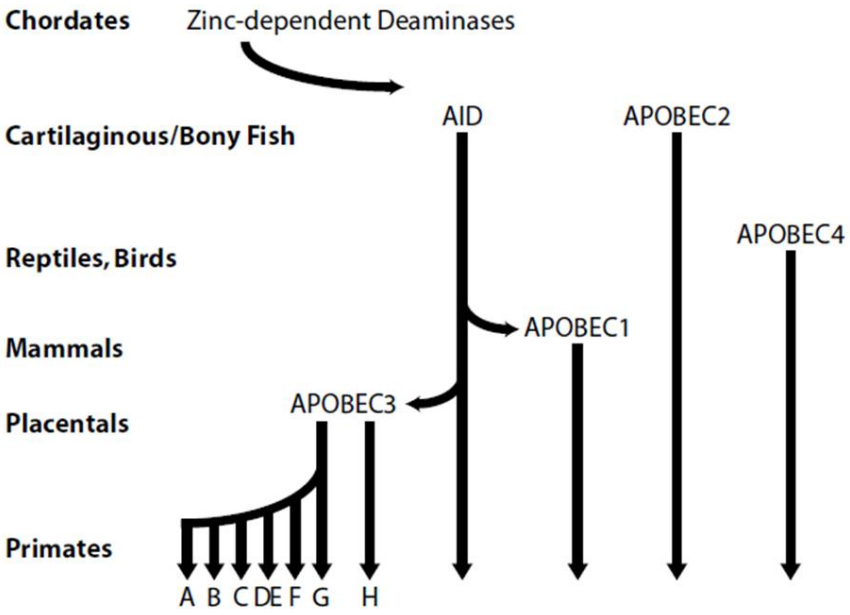
Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3 (APOBEC3) proteins are intrinsic restriction factors of the innate immune system (14). They provide a critical post-entry block against a variety of retroviruses from different species (37). These proteins were originally discovered by Sheehy *et al.* in 2002 during their investigations of the HIV-1 virus (12). They were the first to observe that when the *Vif* gene is removed from HIV-1 (HIV-1Δvif), some cell lines do not allow HIVΔvif viruses to produce infectious particles (12). Their studies concluded that there were cellular factors acting against HIV-1 that could inhibit the infectious cycle of the virus in the absence of the vif protein (12). Subsequent subtractive mRNA hybridization assays revealed that a cellular restriction factor, initially called CEM15, was responsible for the antiretroviral activity (12). Eventually, CEM15 was discovered to be an enzyme with cytidine deaminase activity and due to its structural and functional similarities with other cytidine deaminase enzymes such as APOBEC1 and APOBEC2, it was named Apolipoprotein B mRNA-editing enzyme catalytic polypeptide like-3 G (APOBEC3G) (A3G) protein (Figure-4B) (12). After the initial discovery of A3G, many other family members were identified in humans and other mammalian species (12, 26, 54). The antiretroviral activity of APOBEC3 proteins has been intensely studied over the last nine years. APOBEC3 proteins target single-stranded DNA for their enzymatic activity. They deaminate cytidines to uridines, disrupting the single-stranded DNA sequences and causing in hypermutation of the proviral genome (12, 54). Since the reverse transcription of the viral genomic RNA occurs in the host cell cytoplasm, retroviral DNA intermediates become a major target for APOBEC3 proteins (3, 12, 54).

1.3.2 Cytidine deaminases

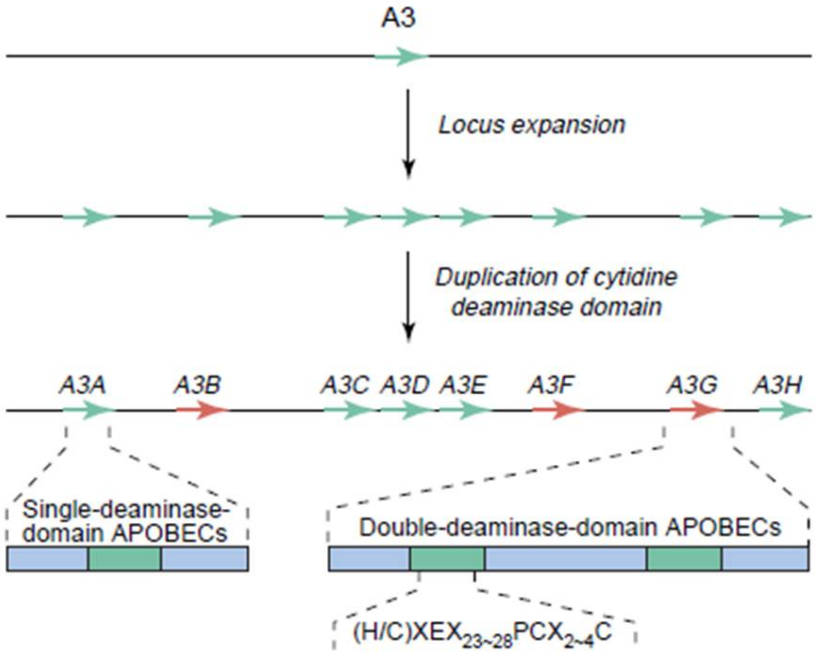
The APOBEC protein family have distinct roles in metabolic events of vertebrate species (9, 12, 15, 54). APOBEC1 (A1) is the only protein in the APOBEC family that targets RNA whereas other members of the family cause mononucleotide-editing in DNA (54). APOBEC proteins have an amino acid motif (H/C)(A/V)E(X_{24-30})(PCXXC) that is conserved among the family members. This signature motif is indispensable for the enzymatic activity, as proven by various mutational studies (3, 11, 12, 42, 49). The deamination event occurs when specific residues at the active site such as histidine (H) and cysteine (C) organize the zinc ion and water molecule to interact with glutamic acid (E). This generates the hydroxyl nucleophile involved in proton shuttling to convert cytidine into uridine (24). The APOBEC family of proteins usually target the carbon atom at the C4 position of cytidine for hydrolytic deamination, which results in the conversion of cytidine to uridine (24, 54). Activation-induced cytidine deaminase (AID) is an important factor for class switching recombination, somatic hypermutation, and gene conversion events (5, 14, 54). AID and APOBEC2 (A2) are considered to be the ancestors of the mononucleotide-editing cytidine deaminase enzymes and AID is thought to be the single ancestor of other family members such as A1 and A3 (Figure-4A) (3, 12, 15, 52, 54). The initial appearance of AID homologues dates back 500 million years in cartilaginous/bony fish and evidence also supports that the earliest appearance of A2 happened approximately during the same period (14). A1 and A3 proteins have evolved from the duplication of AID in placental mammals around 170 million years ago and the subsequent expansion of A3 family proteins then occurred rapidly in primates (14).

Figure-4: A) The evolution of the *APOBEC3* gene family (77). AID and APOBEC2 are considered to be the origins of this gene family (77). B) The evolutionary expansion of human *A3* family genes on chromosome 22 (11). The green arrows on the black lines represent the A3 genes with single deaminase domain and red arrows shows the A3 genes with double deaminase domains (11). The catalytic domain organization of the A3 genes is shown with blue and green boxes and each green box indicates the signature motif of the A3 deaminase domain (11).

A)


Chordates
Zinc-dependent Deaminases
Cartilaginous/Bony Fish
AID
APOBEC2
Reptiles, Birds
APOBEC4
Mammals
APOBEC1
Placentals
APOBEC3
Primates
A B C DE F G H

B)


A3
Locus expansion
Duplication of cytidine deaminase domain
A3A
A3B
A3C
A3D
A3E
A3F
A3G
A3H
Single-deaminase-domain APOBECs
Double-deaminase-domain APOBECs
(H/C)XEX23~28PCX2~4C

1.3.3 APOBEC2 protein

Interestingly, APOBEC2 appears to have only a single catalytic domain, compared to some human APOBEC3 genes such as hA3G, which has two catalytic domains (12, 54). The A2 protein is expressed abundantly in cardiac and skeletal muscle cells, yet it's physiological role in these cells is currently unknown (15, 54). A2 has no apparent catalytic or antiretroviral activity, and oddly animal studies on APOBEC2-deficient mice did not show any abnormalities (12, 15, 52, 54). A2 has the least sequence similarity to AID and APOBEC3 proteins (50% similarity) but it is nevertheless a good model for structural APOBEC3 studies in addition to the fact that it is easily purified and crystallized (15, 52).

1.3.4 APOBEC3 family proteins and their structural organization

After the initial discovery of APOBEC3 proteins, it was revealed that the signature motif of these proteins ((H/C)(A/V)E(X_{24-30})(PCXXC)) was conserved among all family members (54). Each deaminase domain in the A3 structural organization is also called catalytic domain (CD) because the domains in the A3 protein structure contains a signature motif that coordinates the enzymatic activity (6, 11, 23, 26).

Further investigations of APOBEC3 proteins revealed that humans have seven APOBEC3 genes (*A3A, A3B, A3C, A3DE, A3F, A3G* and *A3H*) all of which are located within a 150kb region on chromosome 22 (3, 12, 54, 56). In contrast, the feline genome has four *A3* genes, the equine genome has six, artiodactyls have two or three, and the mouse genome only has a single copy of *A3* gene (37, 64). Structural and sequence analysis of the APOBEC3 proteins indicate that even though they have high sequence similarities, their physiological activity and domain organization may vary in different species (37). It's believed that polymorphic

differences among A3 proteins come from the strong selective pressure in nature and even alternative splicing events that occur in mice, felines and artiodactyls (37, 64). It would be beneficial to the organism to have a variety of A3 proteins to combat a variety of pathogens.

Some human APOBEC3 proteins contain only one catalytic domain (A3A, A3C, and A3H) whereas the other four human A3 proteins (A3B, A3DE, A3G, and A3F) and mouse A3 protein have two catalytic domains (Figure-5A) (3, 12, 26, 36, 54). Each domain has a distinct role in the antiretroviral activity of the protein. Single domain APOBEC3 proteins appear to have evolved before the double domain proteins, indicating that they may be the ancestors of these double domain proteins (3, 12, 37). Also, studies have shown that A3 proteins with two catalytic domains retain elevated antiretroviral activity by using either one of their catalytic domains for RNA binding or for deamination (Figure-5B) (3, 12). The most studied A3 protein is the human APOBEC3G (A3G) protein, and because of its potency as an antiretroviral restriction factor and its utmost importance for HIV-1 infection, the hA3G protein has drawn more attention than any other family member (3, 11, 12). The human A3G protein has two catalytic domains, yet only the C-terminal domain is enzymatically active (12). The N-terminal domain is responsible for RNA binding and viral encapsidation (12). Although the N-terminal domain lacks enzymatic activity, it's believed that RNA binding activity of this domain physically interacts with the reverse transcriptase and inhibits the infectivity of the virus without editing any residues on the single stranded DNA intermediate (3, 11, 12, 15, 26, 43).

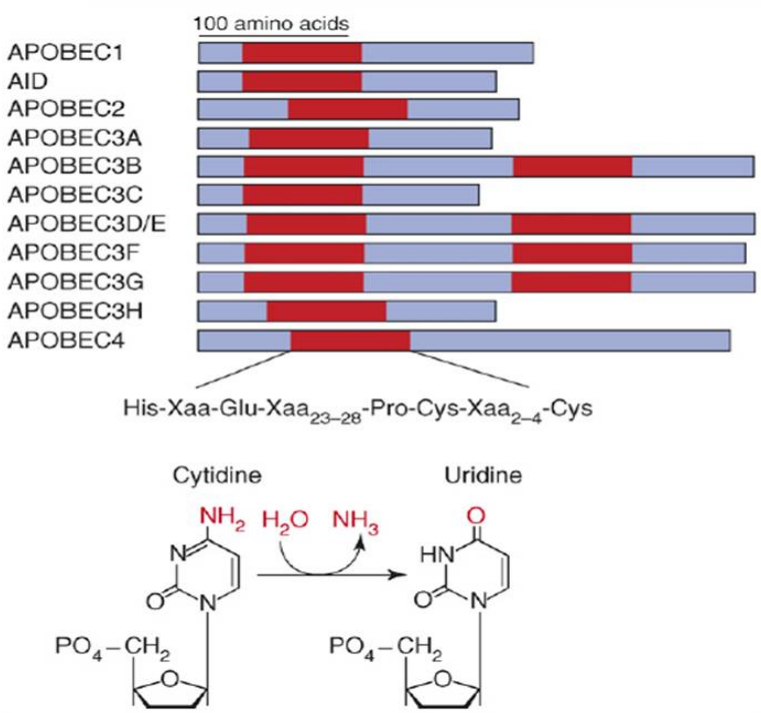
Mice only have one *A3* gene, and unlike its human counterparts, it is located on chromosome 15 (12). Murine A3 has two catalytic domains and the structural organization and functional roles of the catalytic domains are believed to be different than its human counterparts (8, 12,

23, 74). Mouse A3 was reported to have a reversed structural organization compared to human A3G protein (23). The N-terminal domain of the protein has the enzymatic activity responsible for deamination and the C-terminal domain has RNA binding and encapsidation abilities (23). Each of these catalytic domains in both human and mice provide their own contribution to the antiretroviral activity of the A3 protein and this is the reason why double domain APOBEC3 proteins provide enhanced restriction during the viral replication cycle (49, 67, 79).

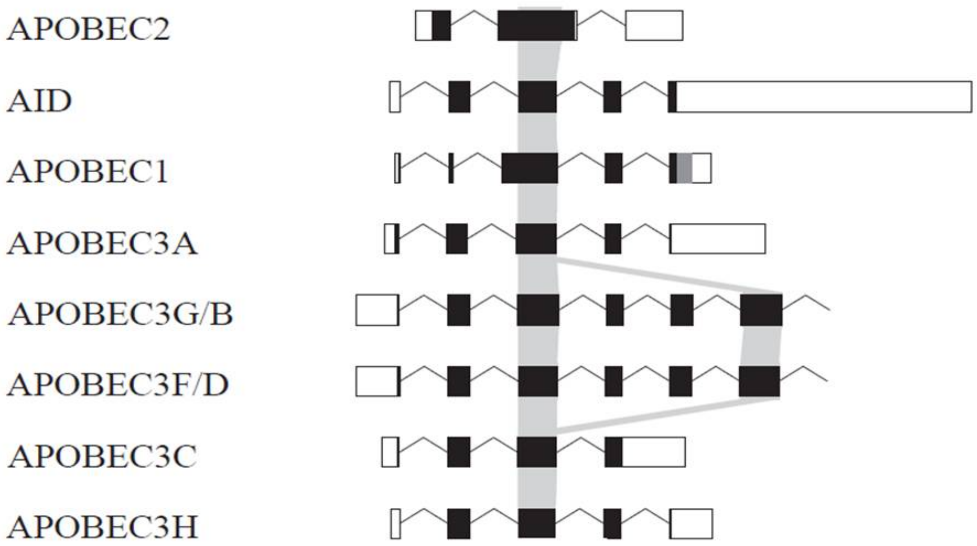
The single *A3* gene in rodents suggests the A3 locus in the human genome, with seven different *A3* genes, branched out after the genetic split between rodents and humans (3, 12, 15, 37). In addition, different *A3* genes in humans diversified as a result of tandem gene duplication and recombination events with uneven crossover (3, 12, 15, 37). Even though murine A3 has reversed structural organization and 30% sequence identity to the human A3G protein, both proteins are highly active against HIVΔvif since the anti-APOBEC3 activity of Vif is absent (19, 57, 78). The mA3 can also inhibit the replication cycle of wild-type HIV-1 because Vif is not able to interact with murine A3 protein (19, 57, 78).

Figure-5: A) Structural organization of the catalytic domains of the human A3 proteins (26). Red boxes indicate the zinc-coordinating motif of the proteins (26). Critical residues in the signature motif coordinate the deamination event and convert cytidines to uridines (26) B) Exon boundaries of APOBEC proteins (51). Filled black boxes represent the translated regions of the proteins and empty boxes shows the untranslated regions (51). The exons that correspond to the signature motif ((H/C)(A/V)E(X_{24-30})(PCXXC)) of the proteins shaded in grey (51). The exons 5, 6, and 7 in APOBEC3B, APOBEC3D, APOBEC3F, and APOBEC3G proteins occurred after gene duplication event of exons 2, 3, and 4 in the same proteins (51).

A)


100 amino acids
APOBEC1
AID
APOBEC2
APOBEC3A
APOBEC3B
APOBEC3C
APOBEC3D/E
APOBEC3F
APOBEC3G
APOBEC3H
APOBEC4
His-Xaa-Glu-Xaa23–28-Pro-Cys-Xaa2–4-Cys
Cytidine
Uridine
NH2
H2O
NH3
O
HN
N
PO4 – CH2

B)


APOBEC2
AID
APOBEC1
APOBEC3A
APOBEC3G/B
APOBEC3F/D
APOBEC3C
APOBEC3H

1.3.5 High and low molecular mass complexes of APOBEC3 proteins

Interestingly, the activity of APOBEC3 proteins is regulated by an unknown mechanism. However, it is clear that A3 enzymatic activity is closely regulated to protect the cell's genomic DNA from APOBEC3 catalytic activity during cell division (12). A3 proteins are generally localized in the cytoplasm; however some family members like hA3B reside in both the cytoplasm and nucleus (12). Depending on the current activation status of the cell, A3 proteins either form 5-15 MDa high molecular mass (HMM) ribonucleoprotein complexes or 150 kDa low molecular mass (LMM) ribonucleoprotein complexes (12). When they are bound into HMM complexes, A3 enzymatic activity is consequently impaired, preventing their DNA deamination activity (12, 47, 70). In contrast, their catalytic activity to edit DNA sequences is not affected when they are in LMM complexes (12, 47, 70). Investigations on the formation of A3 protein complexes indicate that an RNA component is involved in HMM complex formation (12, 47, 70). Therefore, it has been hypothesized that A3 proteins are capable of binding RNA molecules and also each other during the formation of HMM complexes (12, 47, 70).

The involvement of cytokines such as IL-2 and IL-15 in the formation of HMM complexes has been shown by different groups (12). These cytokines can be produced by lymphoid tissues abundantly and are able to provoke the assembly of HMM complexes (12). In addition, transformation of LMM A3 proteins to HMM complexes can be achieved by activating T cells with certain cytokines (IL-2, IL-7, IL-15 and anti-CD3/CD28) (12).

The cytoplasmic localization of hA3G proteins is of utmost importance because of its interactions with other cellular factors and incorporation into budding viral particles, not to

mention that a cytidine deaminase enzyme would be extremely hazardous in the nucleus. Investigations on A3 molecular mass complexes revealed that when A3 proteins assemble to form the HMM complexes, they interact with cellular RNA binding proteins, mRNAs and small non-coding RNAs (47). In addition, biochemical studies on human A3G revealed that miRNA-induced silencing complexes, RoRNPs, processing bodies, stress granules, Staufen granules as well as ribosomes also actively interact with hA3G protein therefore they may have a significant role in the formation of HMM complexes (47). Interestingly, human A3G is thought to localize to lipid rafts, and as a result, this may help its assembly into HMM complexes (47, 55). Experiments aimed to understand the importance of lipid rafts in the assembly of molecular mass complexes categorized human A3G proteins into three groups according to their interactions with lipid rafts: LMM hA3G (soluble A3G), raft-associated LMM hA3G and HMM hA3G complexes (47). Each group was quantified throughout a time course experiment which demonstratedhA3G proteins are in LMM complexes after initial synthesis and then move into lipid rafts and become raft-associated LMM complexes. Upon activation or cell division, these protein complexes then become HMM hA3G complexes (47). Previous studies on hA3G incorporation into HIV-1 budding particles hypothesized that HIV-1 viral particles may only associate with newly synthesized hA3G proteins because they tend to form HMM complexes with RNAs and other unknown factors in the cell after their synthesis. These complexes cannot be encapsidated sufficiently, and only a few studies have supported the idea that HMM hA3G complexes may go into budding particles albeit in very low numbers (12, 47, 70)

In summary, lymphocytes are the main targets for most retroviral pathogens and the expression of A3 proteins in these cells provide a strong intracellular block against retroviral

pathogens (26, 35, 58, 78). Considering the importance of the enzymatic activity of A3 proteins and their formation into different types of molecular mass complexes, the link between these two features may shed light on the restrictive capabilities of the A3 proteins. Initially, hA3G protein complex formation studies were focused on T cells since $CD4^+$ T cells are the main targets for HIV-1 infection, and A3 proteins are abundantly expressed in these cells (11, 12).

Assembly of the molecular mass complexes is not a simple event. There are intricate steps and various cellular factors involved in the formation of these complexes, many of which are currently unknown. Future studies on the identification of crucial hA3G residues responsible for the formation of HMM complexes may provide further insight into these complex cytoplasmic interactions, and may provide insight into the importance of assembly and localization of the A3 proteins.

1.3.6 APOBEC3-mediated restriction of the retroviral replication cycle

Once a retrovirus binds to its specific cell receptor via its Env glycoprotein, the viral envelope fuses to the cell membrane and the core of the retroviral particle is released into the cytoplasm of the target cell. These viral particles then uncoat and the RNA genome undergoes reverse transcription and integration, allowing for expression of viral proteins. After transcription and translation, viral proteins migrate to the cell membrane and assemble into budding viral particles (20, 38). At this stage, APOBEC3 proteins in the cytoplasm bind to the nucleocapsid segment of the viral gag protein and are pulled into budding viral particles. These particles are then released from the infected cell and are free to target other cells (12). When these viral particles attach and release their viral capsid into the new target

cell cytoplasm, APOBEC3 proteins block the viral replication cycle during reverse transcription (12). The antiretroviral activity of A3 proteins takes place in the target cells during the second round of infection. Incorporation of only 7±4 A3 molecules in the viral capsid is sufficient to efficiently hinder the replication of that retroviral particle (11, 12, 26, 79).

In the target cells, the reverse transcription process starts with the elongation of tRNA attached to the viral genomic RNA, allowing for the production of the minus strand strong stop DNA. The following steps involve the translocation of the minus strand stop DNA to the 3' end to generate the whole minus strand (78). Once the minus strand is generated, the RNase H enzyme dissolves the RNA plus strand template and leaves the minus strand DNA intermediate as a single-stranded sequence (78). A3 proteins slide and jump on this single-stranded DNA to deaminate cytidines to uridines, the mutation frequencies of these sequences increase with longer exposure time to A3 proteins. Therefore, the number of mutations are directly related with the exposure time prior to plus strand DNA synthesis (78). When second strand DNA is synthesized, uridine is recognized as thymine and an adenine 'A' is inserted opposite to uridine 'U'. As a result, these mutations in the integrated proviral genomes may cause stop codons or other defects that can affect subsequent transcription and translation events (11, 12, 49, 53, 67).

If the minus-strand DNA intermediate is heavily edited by the A3 proteins, either one of two events can happen. The mutations can either appear on the rapidly synthesized plus-strand DNA and the DNA intermediate can integrate into the target cell genome, or, uracil DNA glycosylase and apurinic-aprimidinic endonuclease enzymes can attack these sequences and induce their degradation in the cytoplasm (11, 12). The regulatory events that decide whether

the mutated DNA sequences will be integrated or destroyed by the cellular enzymes are not yet identified (11, 12, 49, 53, 67).

Unfortunately, some retroviruses such as HIV-1 have evolved distinct mechanisms to avoid the restriction activity of APOBEC3 proteins. HIV-1 encodes a basic 23 kDa viral protein called virion infectivity factor (Vif) that can bind to APOBEC3 proteins, and cause their degradation through the 26S proteasome (12, 48). During the infectious cycle of HIV-1, vif binds the N-terminal region of APOBEC3 and links it to the ubiquitin ligase (E3) complex. This complex recruits several important proteins such as Elongin C, Elongin B, Cullin 5, Nedd8, and Rbx1 (12). This complex will thus promote the polyubiquitylation of APOBEC3 and its subsequent degradation by the proteasome (Figure-1B) (3, 12). As a result of this degradation event, APOBEC3 proteins are excluded from virions and their antiretroviral activity against HIV-1 becomes ineffective (3, 11, 12, 78).

APOBEC3 proteins are believed to exploit two different restriction mechanisms when they interact with retroviral pathogens (26). In the first restriction pathway, APOBEC3 interferes with the retroviral replication cycle when positive-stranded retroviral RNA is reverse transcribed into the minus DNA strand. (12, 18, 34, 36, 49, 74). In the second, APOBEC3 proteins do not use the enzymatic activity of their catalytic domains; yet they still decrease the infectious activity of retroviral pathogens significantly (6, 7, 26). Even though this mechanism is not yet understood, some reports speculate that APOBEC3 proteins may disrupt the activity of reverse transcriptase by affecting its processivity thereby decreasing the accumulation of DNA products (3, 6, 7, 26, 42). It may also impair the interactions between the retroviral RNA and reverse transcriptase, or use other mechanisms that are

currently unknown. (3, 6, 7, 26, 42). Thus, A3 proteins provide a multifunctional restriction mechanism to impede the replication cycle of retroviral pathogens (Figure-6).

1.3.7 The Deamination-dependent restriction pathway of APOBEC3 proteins

A3 proteins from different species have the ability to inhibit the retrovirus replication cycle and the mobility of endogenous retroelements as well as some exogenous retroviruses including HIV-1, SIV and FIV lentiviruses, RSV (alpharetrovirus), MMTV (betaretrovirus), MLV (gammaretroviruses), HTLV (deltaretroviruses) and foamy viruses (12, 26, 37). In addition, Ty1 and Ty2 yeast retrotransposons, MusD, Pmv and IAP mouse endogenous retroelements, and the porcine endogenous retrovirus (PERV) are hindered by A3 proteins (12, 26, 37). Even though the antiviral activity of A3 proteins also has a non-enzymatic component, they usually utilize their enzymatic activity to block the infectious cycle of the viral pathogens through their DNA editing activity. Mutations caused by A3 then get passed to subsequent generations of retroviruses (35, 36, 49, 67).

As previously mentioned, endogenous retroviruses occupy a large portion of the human genome. It has been revealed that many of these inactived endogenous retroviruses contain a substantial mutational burden in their sequences (16, 17, 28, 66). After the discovery of A3 proteins, studies on endogenous retroviruses indicated that A3 proteins may have played a crucial role in inactivating retroviral pathogens by generating C to U mutations on the minus DNA intermediate during reverse transcription (12, 28). Analysis on endogenous murine leukemia viruses detected G-to-A mutations in the integrated genomes of these viruses, the mark of the A3 proteins (28). The enzymatic activity of A3 proteins depends on the context of the sequence as A3 proteins prefer specific motifs for deamination (28). For instance, mA3

protein tends to mutate the third cytidine in the 3' TCC 5'or 3' TTC 5' context (28, 35). This sequence specificity helps to identify the mark of mA3-mediated enzymatic activity. Sequence analysis was completed for endogenous viral sequences to investigate the source of the mutations and to observe whether they were generated by the mA3-mediated enzymatic activity over time (28). In mice, it was shown that there is a considerable amount of these mutations in the genomes of endogenous retroviruses within a TTC or TCC sequence context (28). Also, the gradient of mutation frequency in endogenous MLV genomes was examined to find out whether certain regions in the viral genome contained more mutations than the others. The 3' end of the minus strand DNA stays single-stranded longer than its 5' end during reverse transcription. Due to this exposure time to A3 proteins, the 3' end of the proviral genome has more mutations than the 5' end (28). Sequence analysis on HIVΔvif also indicated that hA3G protein targeted almost 4% of the cytosines in the genome, and that there are hotspots where certain bases are consistently mutated (78). These hotspots may be related to the sequence motif preference of the protein, or other mechanistic factors such as discontinued reverse transcriptase activity (78).

1.3.8 The deamination-independent restriction pathway of APOBEC3 proteins

Since A3 proteins have been discovered, many research groups pointed out the non-enzymatic antiviral activity of A3 proteins. They observed this restriction under different experimental conditions. For example, it has been shown that A3G and A3F can impair the activity of hepatitis B virus without modifying the viral genome with the enzymatic activity of the protein (6). As well, A3A has also a non-enzymatic antiretroviral activity against various retroelements such as intracisternal A, MusD and Line-1 (8). Groups have also found that catalytically inactive A3G proteins still conserve a partial inhibitory effect against HIV-1

(6, 8). Finally, murine A3 proteins can restrict the infectious activity of MMTV, F-MLV and MoMLV in the absence of discernible editing activity (6, 8, 42, 58, 59).

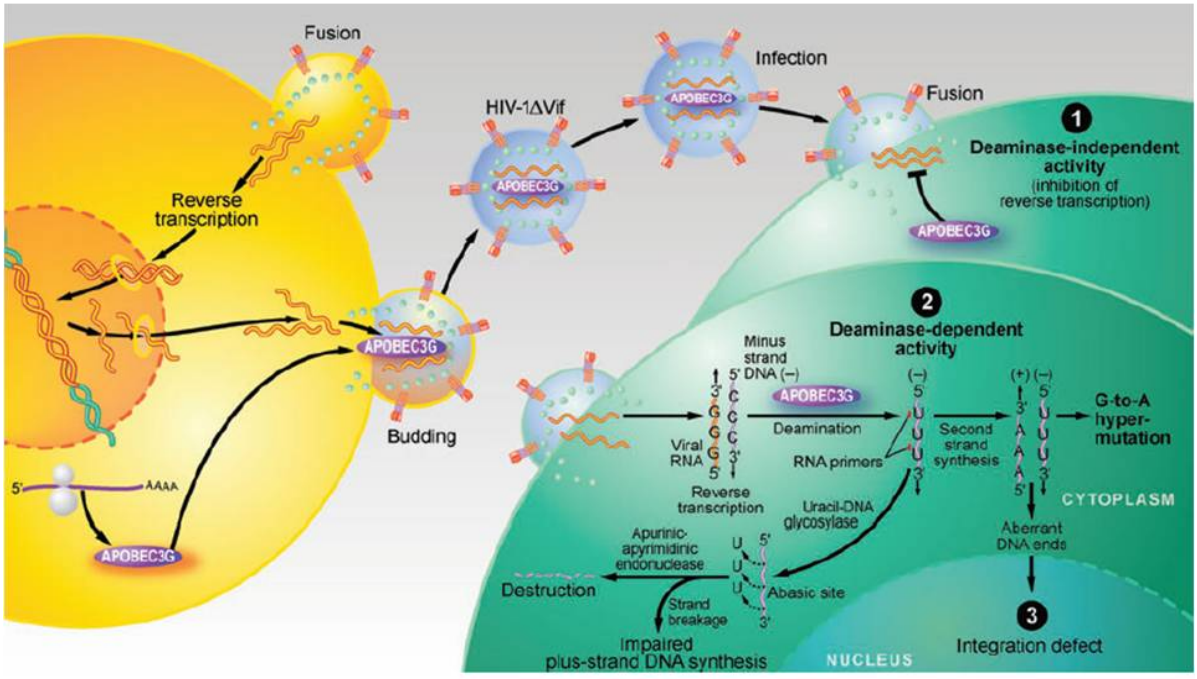
We know that A3 proteins enzymatically edit the viral DNA intermediates during the reverse transcription and sequences that are heavily mutated are presumed to be degraded prior to integration to the target cell genome (6, 12). It is difficult to precisely measure the antiretroviral activity of A3 proteins through the deamination-independent restriction pathway. The number of mutated sequences and the mutation rate induced by A3 proteins may not explain the real capacity of these proteins, and *in vitro* studies may underestimate the restrictive potential of A3 intrinsic restriction factors (6, 7). As a result, further studies are needed to investigate the interactions between A3 proteins and viral RNA, reverse transcriptase, integrase and other mechanisms that play a role in the retroviral replication cycle.

Several groups speculate that A3 proteins may interfere with primer tRNA annealing, minus and plus stand transfer, primer tRNA processivity and removal, DNA elongation as well as proviral integration during the retroviral replication cycle, however the results are controversial (6, 22, 79). For instance, Bishop *et al.* indicated that A3 activity does not interfere with primer tRNA annealing or processing, but inhibits the elongation of reverse transcription products. They speculated that A3 proteins bind to RNA sequences and block the processivity of reverse transcriptase while on the RNA template sequence. (7). Most importantly, the DNA editing activity of the A3 proteins does not correlate with the amount of restriction observed in the studies (6). It was indicated by Bishop *et al.* that there is a definitive relationship between the accumulation of reverse transcripts and inhibition of HIV-1 infection (6). In contrast, Guo *et al*. examined the interactions between the tRNA and

hA3G during reverse transcription and concluded that hA3G impairs tRNA priming, early and late viral DNA production and reduces the infectious activity of retroviral pathogens (22). They also suggested that reduced levels of early DNA transcripts are due to the interaction between tRNA and hA3G, by claiming that hA3G blocks tRNA annealing to viral RNA (22). Although these studies mentioned an array of replication mechanisms that A3 proteins may interact with and inhibit, the existence of A3 proteins in the reverse transcription microenvironment may disrupt all of these mechanisms and slight reductions in each of these steps may generate a strong deamination-independent antiviral activity (7, 26).

Finally, analysis of endogenous levels of A3 proteins and their restriction was crucial for the identification of restriction mechanisms involved in the inhibition of retroviral replication. *In vivo* and *ex vivo* studies have been conducted where primary T cells that contain endogenous levels of A3 proteins were infected with retroviruses. Results indicated that endogenous levels of A3 were able to reduce the infectivity levels of retroviral pathogens; however, the number of detectable hypermutations in the proviral genome was lower compared to *in vitro* tissue culture levels (7, 57, 58). This phenomenon can be attributed to the over expression of proteins in tissue culture assays or other *in vivo* restriction mechanisms yet to be identified.

Figure-6: hA3G-mediated restriction of HIV-1Δvif. hA3G proteins are incorporated into the budding particles after the first round of infection and inhibit the replication cycle of the retroviral pathogens during the second round of infection (12). Restriction of retroviral pathogens is mediated by deamination-dependent and –independent restriction pathways of the protein (12).

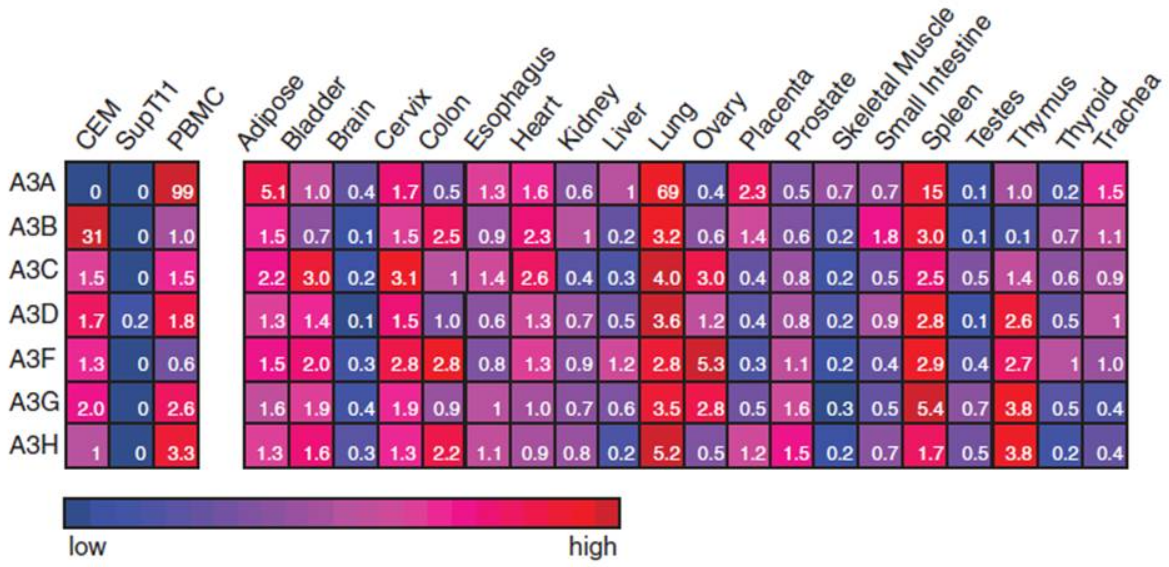

Fusion
Reverse transcription
APOBEC3G
Budding
HIV-1ΔVif
APOBEC3G
Infection
APOBEC3G
Fusion
APOBEC3G
1
Deaminase-independent activity
(inhibition of reverse transcription)
2
Deaminase-dependent activity
Minus strand DNA (–)
APOBEC3G
Deamination
Viral RNA
RNA primers
Second strand synthesis
G-to-A hyper-mutation
Reverse transcription
Uracil-DNA glycosylase
CYTOPLASM
Aberrant DNA ends
Apurinic-apyrimidinic endonuclease
Destruction
Abasic site
Strand breakage
Impaired plus-strand DNA synthesis
3
Integration defect
NUCLEUS

1.3.9 APOBEC3 expression and interspecies restriction

All placental mammals have *APOBEC3* genes but the expression levels of their respective proteins vary among tissues and organs. Distribution and expression levels of the seven A3 proteins vary among all lymphoid and other tissues. APOBEC3 protein expression is crucial in such tissues where the retroviral pathogens hijack the metabolic pathways for their infectious cycle (Figure-7). For instance, peripheral blood lymphocytes are major targets for retroviruses and these cells express substantial levels of A3 and other restriction factors (21). In addition, spleen, thymus, cervix, ovaries and colon are target organs for retroviral pathogens and A3 transcripts of human and mice A3 proteins are expressed abundantly in these tissues (43, 56, 58, 59).

APOBEC3 proteins from different species are able to inhibit the replication cycle of the retroviral pathogens from their own species or cross-species transmission (11, 18, 67, 74). For example, *in vitro* studies on human A3G showed a potent restriction pattern against the Rous sarcoma virus, equine infectious anemia virus (EIAV), feline immunodeficiency virus (FIV), MLV and MMTV. hA3B and hA3C proteins can inhibit the replication of SIV from macaque (SIVmac) and African green monkey (SIVagm), and also, mA3 protein was capable of restricting HIV-1 and primate Foamy Viruses (PFV) (8, 26, 64, 67, 74). The flexible, A3 proteins are also active against certain retroelements like LINE-1, Alu, MusD, HERVs and IAPs that are contained in the human and murine genome (26). Therefore, A3 proteins provide a solid block against intra-species and inter-species transmission of retroviral pathogens (42, 57, 58, 59, 64).

Figure-7: Expression levels of human A3 protein in different tissues (61). Samples were analyzed by qPCR and relative A3 mRNA levels were measured in a variety of cell lines and tissues (61). The numbers in each box shows the relative levels of A3 mRNA from different body parts with the median value in each row set to 1(61).


CEM
SupT11
PBMC
Adipose
Bladder
Brain
Cervix
Colon
Esophagus
Heart
Kidney
Liver
Lung
Ovary
Placenta
Prostate
Skeletal Muscle
Small Intestine
Spleen
Testes
Thymus
Thyroid
Trachea
A3A 0 0 99 5.1 1.0 0.4 1.7 0.5 1.3 1.6 0.6 1 69 0.4 2.3 0.5 0.7 0.7 15 0.1 1.0 0.2 1.5
A3B 31 0 1.0 1.5 0.7 0.1 1.5 2.5 0.9 2.3 1 0.2 3.2 0.6 1.4 0.6 0.2 1.8 3.0 0.1 0.1 0.7 1.1
A3C 1.5 0 1.5 2.2 3.0 0.2 3.1 1 1.4 2.6 0.4 0.3 4.0 3.0 0.4 0.8 0.2 0.5 2.5 0.5 1.4 0.6 0.9
A3D 1.7 0.2 1.8 1.3 1.4 0.1 1.5 1.0 0.6 1.3 0.7 0.5 3.6 1.2 0.4 0.8 0.2 0.9 2.8 0.1 2.6 0.5 1
A3F 1.3 0 0.6 1.5 2.0 0.3 2.8 2.8 0.8 1.3 0.9 1.2 2.8 5.3 0.3 1.1 0.2 0.4 2.9 0.4 2.7 1 1.0
A3G 2.0 0 2.6 1.6 1.9 0.4 1.9 0.9 1 1.0 0.7 0.6 3.5 2.8 0.5 1.6 0.3 0.5 5.4 0.7 3.8 0.5 0.4
A3H 1 0 3.3 1.3 1.6 0.3 1.3 2.2 1.1 0.9 0.8 0.2 5.2 0.5 1.2 1.5 0.2 0.7 1.7 0.5 3.8 0.2 0.4
low
high

1.4 Mouse APOBEC3 protein

Initial studies on A3 were conducted *in vitro,* allowing for different functions of the A3 proteins to be identified in cell culture. However, animal models are required to fully understand the physiological relevance of these functions. To study the basic A3 restriction mechanisms *in vivo*, the mouse can be used as a model. The mouse genome contains only a single copy of the *A3* gene which is located on chromosome 15 (8, 69). Oddly, the *A3* gene knock-out does not cause any growth or reproduction abnormalities or other kinds of immunological defects in these mice (52).

Recent studies indicated that mice and other rodents can be the source of new zoonotic infections and understanding the dynamics of A3 biology could be useful in understanding the zoonotic transmission of retroviral pathogens (39, 56, 60). For instance, the mouse mammary tumor virus (MMTV) and the xenotropic murine leukemia viruses (XMLV) were shown to infect human cells, and their interactions with host A3 proteins play a crucial role their restriction (21, 56, 58, 71). Also, other mouse retroviruses such as Friend murine leukemia virus (F-MLV), Moloney MLV and AKV MLV also interact with mA3 during their replication cycle and investigations on these simple retroviruses can help us better understand the A3 interactions with more complex human retroviruses that deploy advantageous accessory proteins (9, 26, 56, 59, 72, 74). In conclusion, studies on murine A3 protein can broaden our knowledge on the physiological role of the A3 protein family and help generate relevant mouse models that provide mechanistic insight into the restriction of zoonotic transmissions.

In general, sequence analysis of a variety of *Mus* strains and species indicate that mA3 is expressed from two major mouse alleles (35). Studies on inbred mouse strains showed that different strains express either a C57BL/6 or BALB/c mA3 allele (59). The full mRNA transcript of *mA3* contains 9 exons although both major *mA3* alleles also have related mRNA splicing isoforms (35, 59). Different splicing isoforms of the *mA3* gene do not affect the structural organization of the protein and all splicing isoforms contain both catalytic domains (Figure-8A) (8). Although there are different splicing isoforms in various mouse strains, the two major mA3 splicing isoforms either include or exclude exon 5 in their mRNA (Figure-8B) (35, 59, 64). The reason behind this phenomenon is explained by a point mutation in a putative splice branch site sequence and a deleted sequence in the intron between unspliced exon 4 and exon 5 within the mRNA transcripts of the C57BL/6 allele (59). In contrast, the BALB/c allele has also a 7-bp sequence located 53-bp upstream from exon 5 that is almost identical to the canonical branch-site selection sequence, but it is currently unknown whether these sequences contribute to selection of exon 5 in the mRNA transcripts (59). It has been shown that almost all of the mRNA transcripts from the C57BL/6 allele exclude exon 5 (over 95%) (8, 35, 59). The mRNA transcripts from the C57BL/6 allele lacking exon 5 encode a ~49-kDa protein (mA3[C57]ΔE5) that can be found in many inbred mouse strains (35, 59). Also, mRNA transcripts of this allele that include exon 5 are capable of encoding a functional ~51-kDa protein (mA3[C57]), but full-length mRNA transcripts from the C57BL/6 allele constitute less than 5% of the mA3 mRNA pool (35, 59). On the other hand, the BALB/c allele also has the same splicing isoforms that either contain or exclude exon 5, but the distribution of these splicing isoforms are different than the C57BL/6 allele (35, 59). Half the mA3 mRNA from a mouse expressing the BALB/c allele are the full-length mRNA

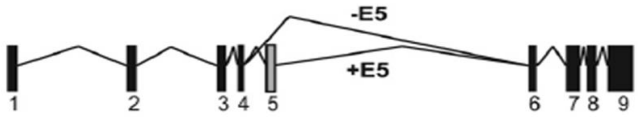
transcripts that encode a ~51-kDa protein and the other half lack exon 5 and encode a ~49-kDa protein (Figure-8B) (35, 59). In addition, some BALB/c mice also express mRNA transcripts that exclude exon 2 and encode a ~38-kDa protein but this splicing isoform can only be detected in very low numbers (35, 59). The full-length mA3 protein of both the C57BL/6 and BALB/c alleles contains 430 amino acids, and the splicing isoform without the exon 5 contains 397 amino acids in both alleles (35, 59, 74). The C57BL/6 and BALB/c alleles have 15 to 17 amino acid differences in their mA3 coding sequences but these polymorphic amino acids have not been associated with antiretroviral activity of the protein (35). In order to confirm the distribution of mA3 splicing isoforms in different mouse strains, splenic RNA was isolated and examined for mA3 mRNA levels by different groups (35, 42, 59, 74). Various mouse strains with the C57BL/6 allele expressed high levels of spliced mA3[C57]ΔE5 mRNA and low levels of full-length mRNA whereas mouse strains with the BALB/c allele expressed high levels of both spliced and full-length mA3[Balb] mRNA in the samples (35, 42, 59, 74).

Figure-8: A) The alignment of mA3 splicing isoforms (35). In the figure, major amino acid changes marked as yellow whereas blue colour amino acids reveal the conservative changes (35). Some residues were underlined to indicate the location of exon 5 and signature motif of the each catalytic domain in the mA3 sequence alignment (35). B) Exon boundaries and alternative splicing of mA3 mRNA (35).

A)

```
                .   .   .   .10  .   .   .   .20  .   .   .   .30  .   .   .   .40  .   .   .   .50  .   .   .   .60
C57_BL6_-E5     MGPFCLGCSHRKCYSPIRNLIS...QETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPV
Balb/c_-E5      MGPFCLGCSHRKCYSPIRNLIS...QETFKFHFKNLRYAIDRKDTFLCYEVTRKDCDSPV
Balb/c_+E5      MGPFCLGCSHRKCYSPIRNLIS...QETFKFHFKNLRYAIDRKDTFLCYEVTRKDCDSPV
Rat             MGPVCLGCSHRRPYSPIRNPLKKLYQQTFYFHFKNVRYAWGRKNNFLCYEVNGMDCALPV

                .   .   .   .70  .   .   .   .80  .   .   .   .90  .   .   100 .   .   110 .   .   120
C57_BL6_-E5     SLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFECAEQIVRFLA
Balb/c_-E5      SLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFECAEQVLRFLA
Balb/c_+E5      SLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFECAEQVLRFLA
Rat             PLRQGVFRKQGHIHAELCFIYWFHDKVLRVLSPMEEFKVTWYMSWSPCSKCAEQVARFLA

                .   .   .   130 .   .   140 .   .   150 .   .   160 .   .   170 .   .   180
C57_BL6_-E5     THHNLSLDIFSSRLYN.VQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRF
Balb/c_-E5      THHNLSLDIFSSRLYN.IRDPENQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRF
Balb/c_+E5      THHNLSLDIFSSRLYN.IRDPENQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRF
Rat             AHRNLSLAIFSSRLYYYLRNPNYQQKLCRLIQEGVHVAAMDLPEFKKCWNKFVDNDGQPF

                .   .   .   190 .   .   200 .   .   210 .   .   220 .   .   230 .   .   240
C57_BL6_-E5     RPWKRLLTNFRYQDSKLQEILR..............................RMDPL
Balb/c_-E5      RPWKKLLTNFRYQDSKLQEILR..............................RVHLL
Balb/c_+E5      RPWKKLLTNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVEGRRVHLL
Rat             RPWMRLRINFSFYDCKLQEIFS..............................RMNLL
                                                    Exon 5
                .   .   .   250 .   .   260 .   .   270 .   .   280 .   .   290 .   .   300
C57_BL6_-E5     SEEEFYSQFYNQRVKHLCY..YHRMKPYLCYQLEQFNGQAPLKGCLLSEKGKQHAEILFL
Balb/c_-E5      SEEEFYSQFYNQRVKHLCY..YHGMKPYLCYQLEQFNGQAPLKGCLLSEKGKQHAEILFL
Balb/c_+E5      SEEEFYSQFYNQRVKHLCY..YHGMKPYLCYQLEQFNGQAPLKGCLLSEKGKQHAEILFL
Rat             REDVFYLQFNNSHRVKPVQNRYYRRKSYLCYQLERANGQEPLKGYLLYKKGEQHVEILFL

                .   .   .   310 .   .   320 .   .   330 .   .   340 .   .   350 .   .   360
C57_BL6_-E5     DKIRSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGL
Balb/c_-E5      DKIRSMELSQVIITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGL
Balb/c_+E5      DKIRSMELSQVIITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGL
Rat             EKMRSMELSQVRITCYLTWSPCPNCARQLAAFKKDHPDLILRIYTSRLYFYWRKKFQKGL

                .   .   .   370 .   .   380 .   .   390 .   .   400 .   .   410 .   .   420
C57_BL6_-E5     CSLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLRRIKESWGLQD
Balb/c_-E5      CSLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLHRIKESWGLQD
Balb/c_+E5      CSLWQSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLHRIKESWGLQD
Rat             CTLWRSGIHVDVMDLPQFADCWTNFVNPQRPFRPWNELEKNSWRIQRRLRRIKESWGL..

                .   .   .   430 .   .
C57_BL6_-E5     LVNDFGNLQLGPPMS
Balb/c_-E5      LVNDFGNLQLGPPMS
Balb/c_+E5      LVNDFGNLQLGPPMS
Rat             ...............
```

B)



1.4.1 Antiviral restriction of the mouse APOBEC3 protein

The mA3 protein is a potent restriction factor against many retroviruses and endogenous retroelements (35, 42, 74). However, some viruses are inhibited partially or not at all by mA3 (57, 58, 67). The antiretroviral activity of mA3 has been shown to be dependent on the type of retrovirus, the mouse strain, and the different alleles and splicing isoforms as well as their expression levels *in vivo* and *in vitro* (18, 35, 42, 56). mA3 provides a powerful barrier against germ line integration resulting from the replication cycle of retroviral pathogens, and protects the host tissues from the pathologic outcome of infection (35, 42, 68, 71, 72).

To investigate the antiviral role of mA3 splicing isoforms Okeoma et al. examined the virion incorporation of these splice variants with the MMTV virus, and also, post-infection levels of MMTV to analyze the restrictive capabilities of the splicing isoforms (59). Analysis indicated that mA3[C57] and mA3[C57]ΔE5 are incorporated into budding viral particles almost equally but post-infection analysis detected fewer infected cells with the mA3[C57]ΔE5 than with mA3[C57] protein isoform (59). Furthermore, *in vitro* studies of murine A3 tested its antiretroviral activity at physiological levels (35, 57, 58, 59). Initially, primary cells were isolated from both C57BL/6 and BALB/c mice and mA3 mRNA expression levels were analyzed. Physiological levels of mA3[C57]ΔE5 and mA3[Balb] were compared in this study and the authors found that mA3[C57]ΔE5 expression was ~5 fold higher than mA3[Balb] mRNA expression (59). These experiments were performed on primary dendritic cells and splenocytes that were isolated from mice with different mA3 alleles (59). In addition, infection assays using primary cells to examine the physiological activity of mA3s also showed less restriction with mA3[Balb] compared to mA3[C57]ΔE5

(59). These results indicate that splicing isoforms and expression levels of these variants in different tissues play a crucial role on the inhibition of retroviral pathogens (35, 42, 58, 59).

MMTV, Friend-MLV and Moloney-MLV viruses are all partially restricted by the murine A3 protein whereas human A3G inhibits the activity of all three viruses (8, 42, 56, 74). In addition, both hA3G and mA3 generate high levels of mutations in the HIVΔvif proviral genome, but the same editing activity for mA3 cannot be detected in the proviral genome of these murine retroviruses (8, 42, 56, 74). Yet, hA3G inhibits every one of these mouse viruses and HIVΔvif to the same level (11). Therefore, murine retroviruses tend to recover from the mA3 restriction or only remain partially restricted in their target tissue (8, 42, 56, 74). It is likely that this is because MMTV, F-MLV and MoMLV do not contain any accessory proteins but exploit distinct mechanisms to intercept the activity of mA3 that are yet to be discovered (8, 11, 26, 27, 79). It is now clear that the restriction of MLV by mA3 is partially due to deamination-independent mechanisms (42, 58, 74). Investigations on both early and late reverse transcription products indicate that mA3 reduces the viral DNA accumulation at early and late stage of infection. This reduction is significant enough to decrease the viral activity to minimum levels (8, 42, 57, 59, 74).

1.5 Relevance of the study to the human health:

Environmental pathogens impose an immense burden to human health, and sometimes, the immune system cannot counteract bacterial or viral pathogens rapidly and prevent the host from infection. Such pathogens with a fast replication cycle are able to circumvent the immune defence mechanisms within a few hours after their initial contact, and become a potential threat to the host species. In recent years, numerous studies on retroviral pathogens have led to the discovery of host-encoded intrinsic restriction mechanisms. Since the discovery of A3 proteins in 2002, many features of this protein family have been revealed, but there is much more to explore. Specifically, previous studies showed that intrinsic immunity is a promising area for the development of novel pharmaceutical therapeutics. Indeed, new treatment strategies are needed to cure major health concerns such as HIV-1. Understanding the structural organizations, functions and evolutionary role of A3 protein could be helpful for pharmaceutical research on drug discovery for HIV.

The A3 field has expanded, and allowed animal models to test our current *in vitro* knowledge. However, without understanding the basic features of these proteins, it is unlikely that further analysis of these animal models will be productive. Understanding mA3 function is important for establishing the foundations of an *in vivo* mouse study and our tissue culture experiments will be highly valuable to improving our understanding between virus and target cell interactions. These examinations are required to determine correct strategies to conduct *in vivo* experiments and for the development of new murine models to study the retrovirus infection patterns on these transgenic knock in/out animal models.

Most of our food supply is supported by an industry that is in a direct contact with creatures in nature. Mice are one of these creatures and they interact with our raw food supplies directly or indirectly. Mice may therefore be or become a vector for the transmission of zoonotic diseases. We also know that rodents have a notorious history of spreading pathogens and causing epidemics and pandemics. Therefore, mouse retroviruses can also be a threat for human health; therefore the study of mA3 can provide useful material to understand how zoonotic transmission of retroviral pathogens can be prevented.

In conclusion, investigating the physiological role of murine APOBEC3 protein as an antiretroviral restriction factor can contribute to current *in vitro* experiments. These murine A3 studies can also establish the foundations of future mouse models to examine retroviral infection patterns and for the development of new pharmaceutical drugs. In addition, zoonotic infections of mouse origin may be blocked by the antiviral activity of A3; therefore *in vivo* mouse studies can certainly be a model for us to learn how to prevent these infections.

2.0 HYPOTHESIS:

Previous reports have described that mA3 possesses either deaminase-dependent or deaminase-independent restriction capabilities against human and mouse retroviruses. We hypothesize that these two pathways for restriction act in concert to restrict retroviral replication and not independently to each other. We also hypothesize that catalytic activity of mA3 is essential to provide permanent retroviral inactivation in light of the observation that retroviral species (all of murine origin) shown to be restricted in absence of deamination are nonetheless infectious to mice at various degrees.

3.0 OBJECTIVES:

To analyze the structural domain organization and retroviral restriction mechanisms of mA3 by:

1. Assessing the antiretroviral activity of the mA3 protein against HIV and MLVs, and to evaluate the cytidine deaminase activity of mA3's two active sites.
2. Determining the significance and the contribution of the deamination-dependent and –independent restriction pathways of human A3G protein and mA3 *in vitro.*
3. Examining the potential role of the glycosylated gag protein of certain gammaretroviruses as a counter mechanism for the antiretroviral activity of mA3.

4.0 MATERIALS AND METHODS

4.1 Cells and expression vectors

Human embryonic kidney epithelium 293T cells and mouse embryonic fibroblast NIH 3T3 cells were cultured in HyClone DMEM/High Glucose medium (Thermo Fischer Scientific, Waltham, Massachusetts, USA) supplemented with 10% fetal bovine serum (FBS) (Sigma-Aldrich, St. Louis, Missouri), 100 U/ml penicillin, 100 µg/ml streptomycin (Multicell, Wisent Inc., Canada) and propagated at 37 ºC in a 5% CO_2 incubator.

Unless otherwise noted, the expression vector encoding replicative Moloney MLV (pMOV-eGFP) virus was identical to the previously described vector (Sliva *et al*., 2004). The green fluorescence protein (eGFP) coding sequence was inserted in the proline-rich region of the *Env* gene of the MoMLV virus (Sliva *et al*., 2004). The viral plasmid constructs encoding replicative AKV-N (AkvU3-eGFP) and AKV-NB (AkvNBU3-eGFP) gammaretroviruses have been detailed by Aagaard *et al.*, 2002 (1). The AKV encoding expression vectors contain an internal ribosome entry site (IRES)-enhanced eGFP protein cassette placed downstream of the U3 region (1, 35). The plasmids encoding replicative Hybrid MLV gammaretroviruses used in this study were generated from MoMLV and AKV-NB constructs and they were initially obtained from the M.S. Neuberger laboratory (University of Cambridge, Cambridge, UK). The expression vector encoding replicative Hybrid-1 MLV (pMOV-AKV hybrid-U5) has the MoMLV backbone but contains the AKV-NB 5'-U5 region and genomic sequence up to the AUG gPr65 Gag initiation codon. The replicative Hybrid-2 MLV (pMOV-AKV hybrid XhoI-BclI) expression vector also has the MoMLV backbone, and the region containing the three glycosylation sites in between the XhoI and

BclI restriction sites have been swapped with a portion of the AKV-NB virus. The single-cycle pseudoviruses, M5P MLV and HIVΔvif, were generated by multi-plasmid expression system using a shuttle vector for MLV and HIV-1, a Gag-Pol expression vector, and a plasmid encoding the VSV-G (Vesicular stomatitis virus G glycoprotein) envelope glycoprotein. HIV-1 shuttling vector is a self-inactivating and packageable plasmid containing an eGFP reporter gene and is expressed from a SFFV (Spleen focus forming virus) 5'LTR promoter (36). The plasmid encoding the HIV-1 Gag-Pol lacking vif (p8.9) has been previously described by Naldini *et al*., 1996. MLV self-inactivating packageable vector M5P MLV, Gag-Pol expression vector, and the vector encoding envelope glycoprotein VSV-G were also obtained from the M.S. Neuberger laboratory (University of Cambrigde, Cambridge, UK) and they have been previously described by Harris *et al*., 2003 and Conticello *et al*., 2003. These plasmids above express viral proteins and RNA genome under the control of a CMV (Cytomegalovirus) promoter (36).

The plasmids encoding human APOBEC3s containing the eGFP reporter gene was detailed by Conticello *et al*. 2003. The eGFP cassette has been removed from the pEGFP-C3 plasmid (Invitrogen Cop., Carlsbad, California, USA) and a Flag sequence was inserted at the NheI and ScaI restriction sites to generate a Flag-tag expression vector (pFlag-C3). APOBEC3 coding sequence was subcloned into the Flag-tag expression vectors at the XhoI and PstI sites to obtain the plasmids encoding Flag-tagged APOBEC3 sequences (36). The hA3G-E69A/E259A mutant was generated from wild-type hA3G expression vector by using the QuickChange XL site-directed mutagenesis kit (Agilent Technologies, Santa Clara, California, USA) according to manufacturer's instruction manual. The mouse APOBEC3 expression vectors and primer sequences have been previously described (35, 74). Total

RNA was isolated from spleens of a variety of inbred mouse strains and primary cDNA clones of mouse APOBEC3 alleles were generated from these samples (35). XhoI and KpnI sites in the pFlag-C3 plasmid have been used to obtain the final PCR products (35). The mA3[C57]ΔE5 plasmid was initially used to generate the catalytic domain point mutants by using site-directed mutagenesis, primer sequences have been previously explained (5, 74). Primers have been used to generate the mA3 SWS mutants in the first and second catalytic domains within the point mutants. The primers for the mA3[C57]ΔE5-SWS-1 mutant: forward, 5'-ACCTGGTATATGTCCGCGAGCCCCTGTTTCGAA-3'; reverse, 5'-TTCGAA ACAGGGGCTCGCGGACATATACCAGG-3'. The primers for the mA3[C57]ΔE5-SWS-2 mutant: forward, 5'-ACCTGCTACCTCACCGCGAGCCCCTGCCCAAAC-3'; reverse, 5'-GTTTGGGCAGGGGCTCGCGGTGAGGTAGCAGGT-3'.

4.2 *In vitro* transfection and infection assays

Adherent 293T cells were used in all transfection assays to produce A3 proteins and infectious viral particles. The day before transfection, 2 x 10^5 cells were seeded per well in a 6-well plate, and incubated overnight. Transfection assays were performed when the cells reached 40% confluence. The growth medium was changed prior to transfection. Expression vectors were mixed with the Genejuice transfection agent (Novagen-EMD4 Biosciences, Darmstadt, Germany), and the ratio of transfection agent to plasmid DNA was 3:1 (3µl Genejuice:1µg plasmid). After transfection, the 293T cells were overlaid with 3 ml of complete DMEM and incubated for 48 hours in order to produce adequate amounts of A3 proteins and infectious viral particles for subsequent assays.

The day before viral infection, 2 x 10^5 NIH 3T3 cells were seeded per well using 6-well plates infected when the cells reached 40% confluence. At 48 hours post-transfection, supernatants containing infectious viral particles (with or without Flag-APOBEC3 expressing plasmids) were collected and filtered through 0.45-µm pore cartridge filters. The NIH 3T3 monolayer was then infected by 0.3 ml virus-containing supernatant in the presence of 8µl/ml of polybrene, and spinoculation was performed at 2000 rpm for 1 hour. Target NIH 3T3 cells, overlaid with 3 ml of complete DMEM and incubated for 24 hours post-infection, were then analyzed by flow cytometry (FACS) (Cyan Cytometer, Beckman Coulter Inc., Brea, California, USA). Expression of eGFP was used to determine the efficiency of infection and/or restriction. For post-infection incubations longer than 24 hours, overlay medium was washed 24 hours post-infection and 3ml of fresh complete DMEM medium was added to each well.

4.3 Virus titration analysis

Transfection and infection assays with viral stocks were performed as previously described in section 4.2. Titration analysis of viral vectors has been shown previously (78). Virus-containing supernatants were filtered through 0.45-µm pore cartridge filters prior to infection, and tenfold serial dilutions were performed on supernatants of each viral stock. Briefly, 2 x 10^5 NIH 3T3 cells were seeded per well in a 6 well plate 24 hours prior to infection. Serial dilutions in triplicates of the virus-containing supernatants were mixed with 300 µl of fresh medium and were added to NIH 3T3 cells in the presence of 8 µl/ml polybrene. The final amount of viral particles in each dilution corresponded to 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, and 10^{-5} µl that was added to the cells. Spinoculation was performed at 2000 rpm for 1 hour and infected NIH 3T3 cells were incubated for 24 hours. Flow cytometry analysis

was used to measure the titer of viral stocks carrying the eGFP reporter gene. Each eGFP positive cell was considered a colony, and dilutions yielding eGFP positive cells in the cell population was used to calculate the titrations with the following equation:

$$\text{Titer (transducing units/ml)} = \frac{\text{Number of target cells (20,000 cells per sample)} \times \left(\frac{\text{\% of eGFP-positive cells}}{100} \right)}{\text{Volume of vector in (ml)}}$$

4.4 Western-blotting analysis

For protein analysis, 2 x 10^5 293T cells were seeded per well in a 6 well plate 24 hours prior to transfection. The plasmids expressing Flag-APOBEC3 were transfected into 293T cells as previously described and incubated for 48 hours. The transfected cell monolayers were then washed with phosphate-buffered saline (PBS) and proteins were extracted with NP40 lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1% NP40, 0.1% sodium deoxycholate, and dH_2O) supplemented with 2X protease inhibitor coctail (Roche, Mannheim, Germany). Cell lysates were mixed with 5X Laemmli buffer (0.5M Tris pH 6.8, glycerol, 20% SDS, bromophenol blue, and dH_2O) and subjected to SDS (sodium dodecyl sulphate) gel electrophoresis using 10% acrylamide gels to separate proteins. Samples were then transferred to 0.45μm pore size PVDF membranes (Millipore, Billerica, Massachusetts, USA) at 0.35 Amp for 1 hour. The membrane was blocked with 5% skim milk in PBS-T (0. 1% Tweener 20) (Acros, New Jersey, USA) for 1 hour at room temperature, or overnight at 4ºC. Membranes were probed with monoclonal antibodies that are reactive with the Flag epitope and β-Tubulin at room temperature for 1 hour, or overnight at 4ºC. Flag epitope tags were visualized by using an Anti-Flag M2-peroxidase HRP-conjugated antibody at a 1:3000

dilution (Sigma-Aldrich, St. Louis, Missouri) whereas β-Tubulin was detected using a 1:5000 dilution of a β-Tubulin HRP-conjugated antibody (1mg/ml) (Abcam, Cambridge, Massachusetts, USA). The antibodies were diluted in fresh 5% skim milk in PBS-T and kept at 4ºC. The membranes were washed three times in PBS-T for 10 minutes each and one time with PBS for 5 minutes before the proteins were detected using the ECL Plus Western Blotting Detection System kit (GE Healthcare, Buckinghamshire, UK), and finally exposed to Clear Blue CL-X Posure imaging films (Thermo Scientific, Rockford, Illinois, USA).

4.5 Sucrose gradient and velocity sedimentation assays

Transfection and protein extraction assays have been performed as previously described in section 4.2 and 4.4. Step gradient solutions were prepared using various concentrations of sucrose (5%, 10%, 20%, 30%, 40%) mixed with sucrose gradient buffer (0.01M Tris-HCl pH 7.4, 0.25M KCl, and 0.01M $MgCl_2$). This gradient was prepared in advance by layering 2.3 ml of 40% sucrose gradient from the bottom to 5% sucrose step gradient at the top in a 13.2 ml polyallomer tube (Beckman Coulter Inc., Brea, California, USA). After each step of layering, sucrose gradients were snap frozen in liquid N_2. Cell extracts were then overlaid onto the sucrose gradient and tubes were gently placed into a black titanium bucket assembly and capped tightly (Beckman Coulter Inc., Brea, California, USA). The SW 41 Ti (6 X 13.2) ultracentrifugation rotor (Beckman Coulter Inc., Brea, California, USA) was used at 41 000 rpm for 6 hours at 4ºC and samples were collected in nine fractions (1.4 ml) starting from the top, after which each fraction was subjected to Western blotting analysis.

4.6 Viral Sequence Analysis

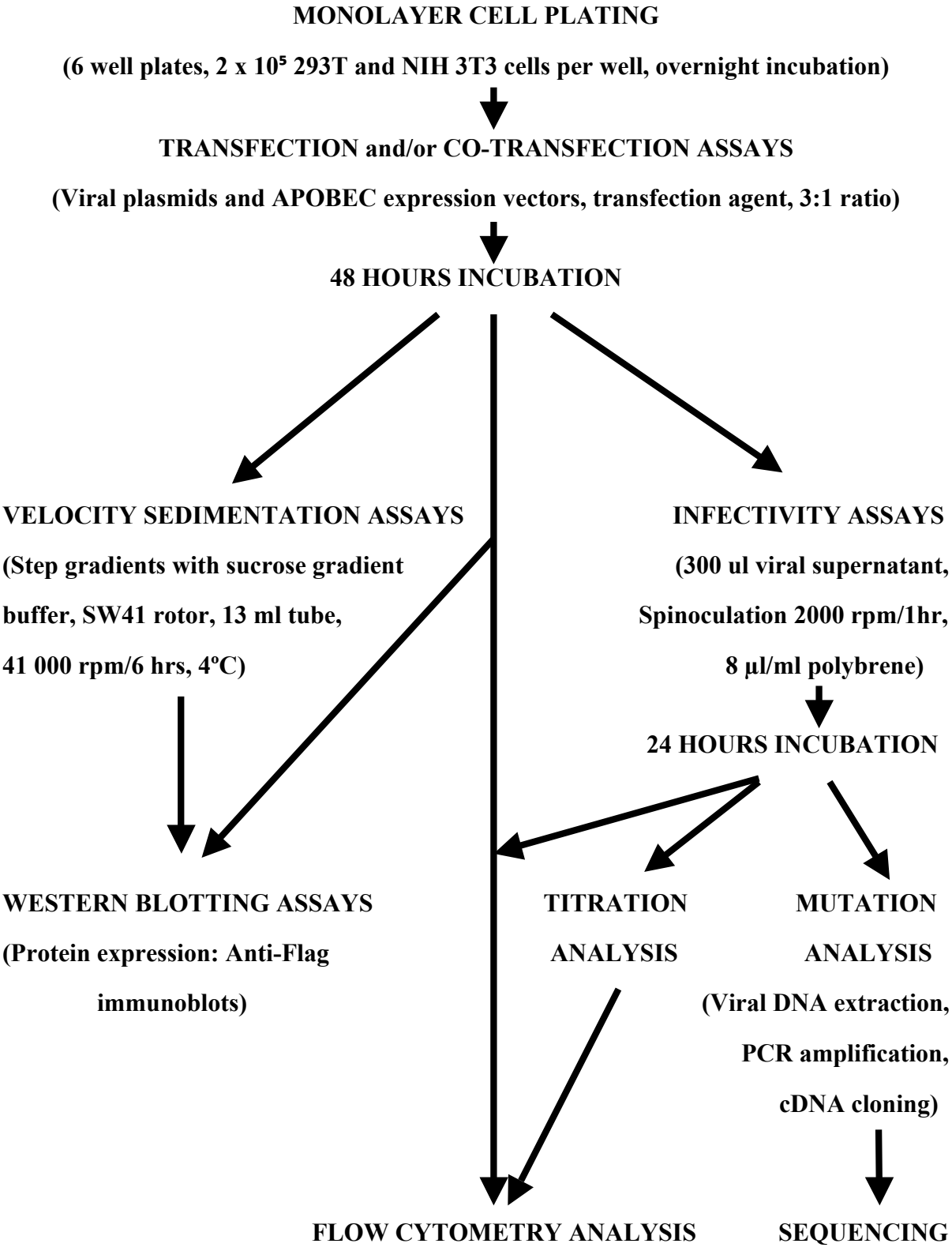
Co-transfection assays were used to produce infectious viral particles with or without the APOBEC proteins and infection assays were performed on monolayer NIH 3T3 target cells (refer to section 4.2 in materials and methods for complete protocol). Twenty four hours post-infection, genomic DNA of the NIH 3T3 cells was extracted using the Wizard genomic DNA purification kit (Promega, Madison, Wisconsin, USA) according to manufacturer's protocol sheet. Samples were then treated with the DpnI restriction enzyme for 1 hour at 37ºC to remove the plasmid carryover from the co-transfection assays. DpnI restriction enzyme was denatured at 80ºC for 20 minutes, and then samples were incubated at room temperature for 5 minutes. The high fidelity polymerase Prime Star (Takara Bio Inc., Otsu, Shiga, Japan) was used to amplify the viral DNA of HIVΔvif, MoMLV, Hybrid-1 MLV, Hybrid-2 MLV, and AKV-NB viruses from infected NIH 3T3 cells. The following primers were selected to target the 717-bp segment of the eGFP coding sequence: forward, 5'-GTGAGCAAGGGCGAGGAGCTGTTCA-3'; reverse, 5'-CTTGTACAGCTCGTCCATGC CGAGA-3' and the amplification reaction was performed as follows: total reaction volume, 50 µl; template concentration, 30-50 ng of total cellular DNA/reaction, 10 µl 5X Taq buffer, 4 µl 10µM dNTPs, 1µl of forward primer, 1 µl of reverse primer, and 0.5 µl Primer Star Taq polymerase (Takara Bio Inc.). The PCR program was 40 seconds at 94ºC, (10 seconds at 94ºC, 10 seconds at 55ºC, 60 seconds at 72ºC) X 32 cycles and 1 minute at 72ºC. The correct band size was confirmed with 1% agarose gel electrophoresis and samples were gel purified using the Wizard SV gel and PCR clean-up system kit (Promega, Madison, Wisconsin, USA). Gel purified samples were then diluted to 20 ng/ml and cloned using the Zero Blunt TOPO PCR cloning kit (Invitrogen Corp., Carlsbad, California, USA). The standard M13

reverse primer was used for sequencing, and positive samples were analyzed utilizing Sequencher DNA sequencing software (Gene Codes Cop., Ann Arbor, Michigan, USA). Spurious sequences were expelled from the dataset prior to alignment and mutation analysis, and confirmed viral sequences were scrutinized in order to calculate the mutation rate and mutation specificity for each sample.

4.7 Statistical Analysis

All data from stability, restriction and infectivity assays (n=3) excluding figures-37, -38, -39 (n=2) are plotted with standard error of the mean (SE).

Figure-9: Schematic description of experimental model used in the study.


MONOLAYER CELL PLATING
(6 well plates, 2 x 10⁵ 293T and NIH 3T3 cells per well, overnight incubation)
TRANSFECTION and/or CO-TRANSFECTION ASSAYS
(Viral plasmids and APOBEC expression vectors, transfection agent, 3:1 ratio)
48 HOURS INCUBATION
VELOCITY SEDIMENTATION ASSAYS
(Step gradients with sucrose gradient buffer, SW41 rotor, 13 ml tube, 41 000 rpm/6 hrs, 4ºC)
INFECTIVITY ASSAYS
(300 ul viral supernatant, Spinoculation 2000 rpm/1hr, 8 µl/ml polybrene)
24 HOURS INCUBATION
WESTERN BLOTTING ASSAYS
(Protein expression: Anti-Flag immunoblots)
TITRATION ANALYSIS
MUTATION ANALYSIS
(Viral DNA extraction, PCR amplification, cDNA cloning)
FLOW CYTOMETRY ANALYSIS
SEQUENCING

5.0 RESULTS

5.1 Assessment of the biology of the mA3 splicing isoforms and the retroviruses

5.1.1 Establishing an *in vitro* model to analyze the antiretroviral function of the mA3 protein

To examine the antiretroviral role of the mA3 splicing isoforms in tissue culture, an *in vitro* model was required to perform a proper assessment. This model was designed to investigate the antiviral activity of the mA3 protein after single-cycle retroviral infection, since its main role within the innate immune system is to provide an early post entry block against retroviral pathogens. It is important to mention that *in vitro* systems may only be partially representative of the physiological function of mA3 in the body.

5.1.2 Expression and protein stability analysis of the mA3 protein splicing isoforms

5.1.2.1 Expression levels and stability of the mA3 splicing isoforms using an eGFP reporter gene

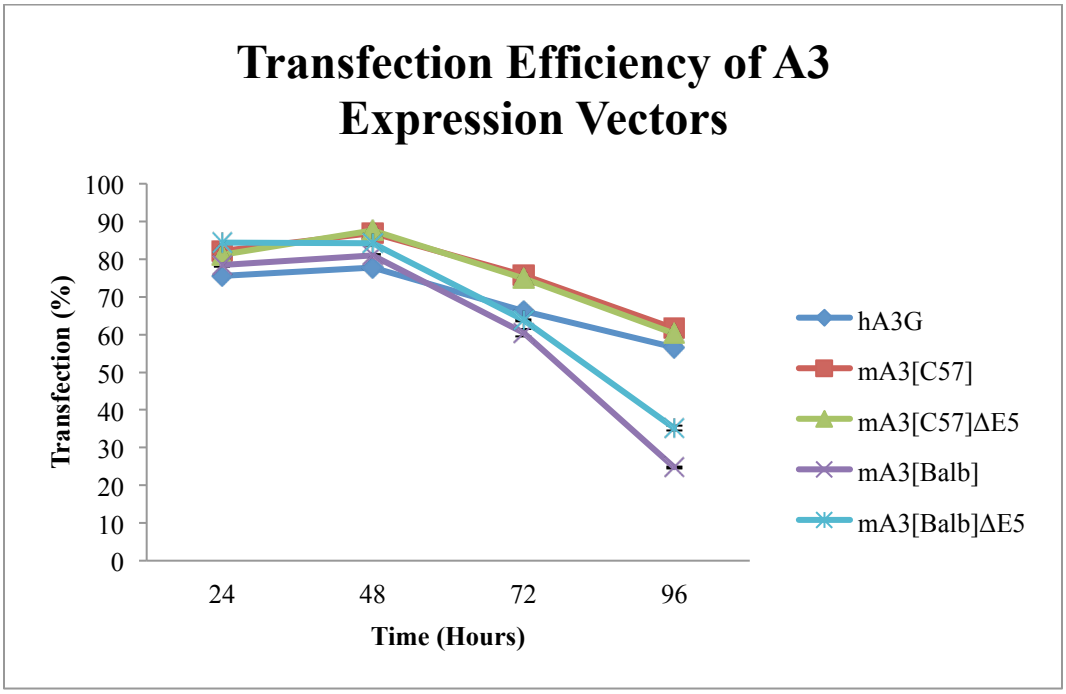
In order to quantify expression, mA3 splicing isoforms and hA3G have been tagged with an eGFP reporter gene. Expression levels as well as the kinetics of plasmid degradation were measured in time course experiments. Simple transfection assays were performed in 293T cells in 10cm petri dishes, incubated for 24 hours, and then trypsinized and prepared for flow cytometry analysis. The remaining transfected cells that were not analyzed, were replated again using 6 well plates allowing for examination at later time points. Flow cytometry analysis was used to quantify the percentage of transfected cells (Figure-10A) by measuring eGFP fluorescence. The intensity of the expression levels was also analyzed for each

transfection by measuring the mean fluorescence (FL-1 channel) at various time points (24, 48, 72, 96 hours) (Figure-10B). It is important to note that more than one mA3 expression plasmid may enter the cell by transfection; therefore mean fluorescence of each cell will reflect this phenomenon.

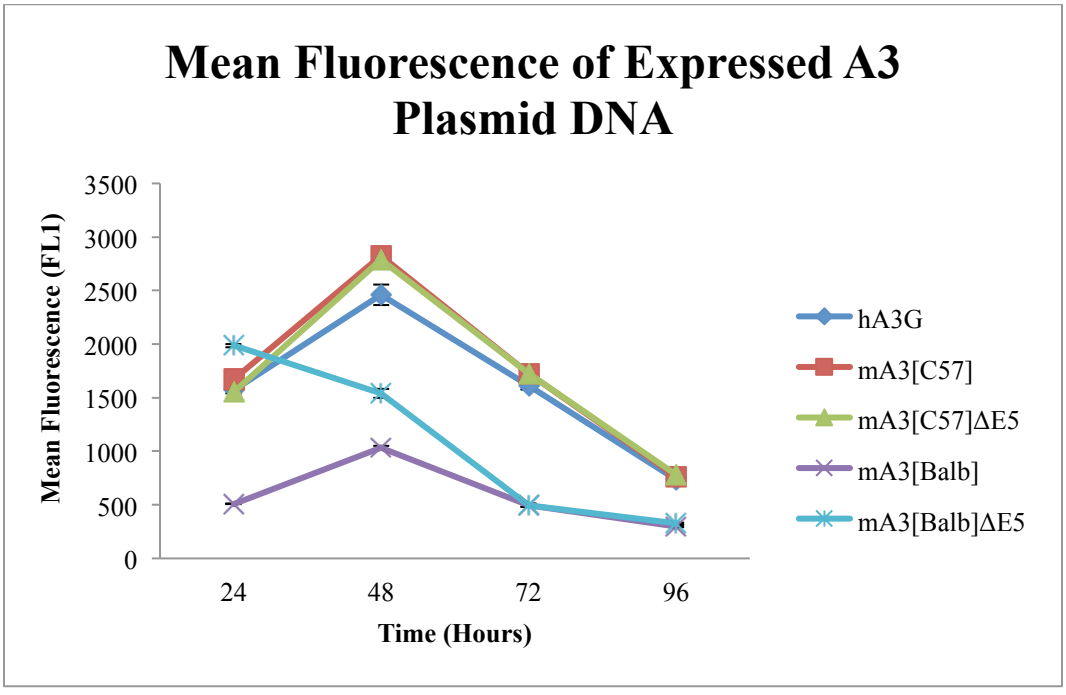
The results indicate that 48 hours post-transfection yields the maximum number of transfected 293T cells as well as peak mean fluorescence measurements for the four mA3 isoforms tested (Figure-10A, -10B). Even though the transfection efficiency was similar for the mA3 isoforms and hA3G at the 48-hour time point (Figure-10A), the mean fluorescence analysis revealed considerable variation among the samples at the same time point (48 hours) (Figure-10B). The fluorescence intensity of hA3G, mA3[C57], and mA3[C57]ΔE5 expression vectors were higher than mA3[Balb] and mA3[Balb]ΔE5 plasmids after 48 hours of incubation. Moreover, mA3[Balb] expression vector had the lowest mean fluorescence (Figure-10B) and transfection efficiency, a pattern which was seen in subsequent assays. This approach revealed the optimal time at which we should perform our experiments. These expression experiments also provided information concerning the depletion of these expression vectors over time *in vitro.*

Figure-10: mA3 protein splicing isoforms were tagged with an eGFP reporter gene and their expression levels were quantified with the CyAn flow cytometer. A) mA3 splicing isoforms and hA3G expression vectors (1 μg) were transfected in 293T cells and the percentage of green fluorescent cells were measured over four consecutive days. Any cell that contains an eGFP tagged A3 protein will be eGFP positive and Transfection (%) represents eGFP positive cells. B) Mean fluorescence levels (FL-1 channel) were also examined to determine the expression and the protein degradation analysis of the mA3 splicing isoforms.

A)


Transfection Efficiency of A3 Expression Vectors
Transfection (%)
100
90
80
70
60
50
40
30
20
10
0
24
48
72
96
Time (Hours)
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]
mA3[Balb]ΔE5

B)


Mean Fluorescence of Expressed A3 Plasmid DNA
Mean Fluorescence (FL1)
3500
3000
2500
2000
1500
1000
500
0
24
48
72
96
Time (Hours)
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]
mA3[Balb]ΔE5

5.1.2.2 Investigating the expression of mA3 isoforms by western immunoblotting

To further investigate the *in vitro* expression of the *mA3* isoforms, these genes were cloned downstream of a Flag epitope tag and their expression levels were measured using an anti-Flag monoclonal antibody in a western immunoblot (WB). Various mA3[C57]ΔE5 point mutants were also analyzed.

The Flag-mA3 isoforms (0.5 μg) were transfected into 293T cells in 6-well plates and incubated for 48 hours. After 48 hours, the transfected cells were collected in cold PBS and lysed with 700 μl of NP40 lysis buffer containing 2X protease inhibitor cocktail. In view of the low expression levels of the mA3[Balb] protein in the previous experiment (Figure-10B), three times more plasmid DNA of the mA3[Balb] (1.5 μg) was transfected in order to increase the expression levels of this protein to levels similar to the other splicing isoforms (Figure-11). Low expression levels of mA3[Balb] protein was also shown by other studies and it was concluded that our expression vectors with eGFP and Flag epitope tag are not the reason behind this phenomena.

In addition, point mutants of the mA3[C57]ΔE5 protein were generated in order to investigate the structural organization of the mA3 protein and the functional role of the catalytic domains (Table-2). The point mutants of the mA3[C57]ΔE5 protein also had the Flag epitope tag in their constructs. Wild-type mA3[C57]ΔE5 and the five point mutants were transfected into 293T cells (1μg), followed by a 48 hour incubation period. The cells were then harvested and lysed as described above, and expression levels were detected by WB (Figure-12). The five point mutants of mA3[C57]ΔE5 protein showed similar expression levels to the wild-type protein, therefore it was concluded that the point mutants generated by

site-directed mutagenesis did not affect the expression levels of these proteins (Figure-12). Also, there was no direct evidence that the eGFP reporter gene nor the Flag epitope tag interfered with the expression levels of the mA3 splicing isoforms or the point mutants of the mA3[C57]ΔE5 protein. However, because all the retroviruses used in this study express the eGFP reporter gene, it was decided to continue with the Flag-tagged mA3 protein in order to analyze the protein expression levels by WB and to measure the infectivity of the retroviruses by the flow cytometry analysis.

Figure-11: The mA3 protein splicing isoforms were tagged with the Flag epitope and expression levels were measured with an anti-Flag western immunoblot. The mA3 splicing isoforms were transfected into 293T cells and the samples were incubated for 48 hours to reach the optimum expression levels. Three times more plasmid DNA was used for the mA3[Balb] splicing isoform to obtain the similar expression levels among the mA3 splicing isoforms.

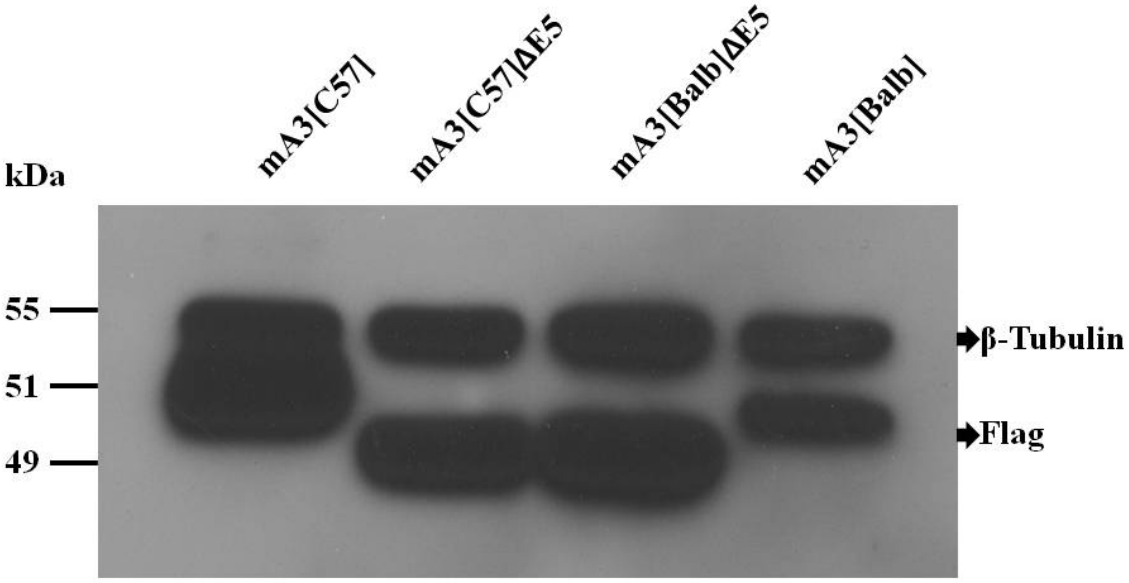
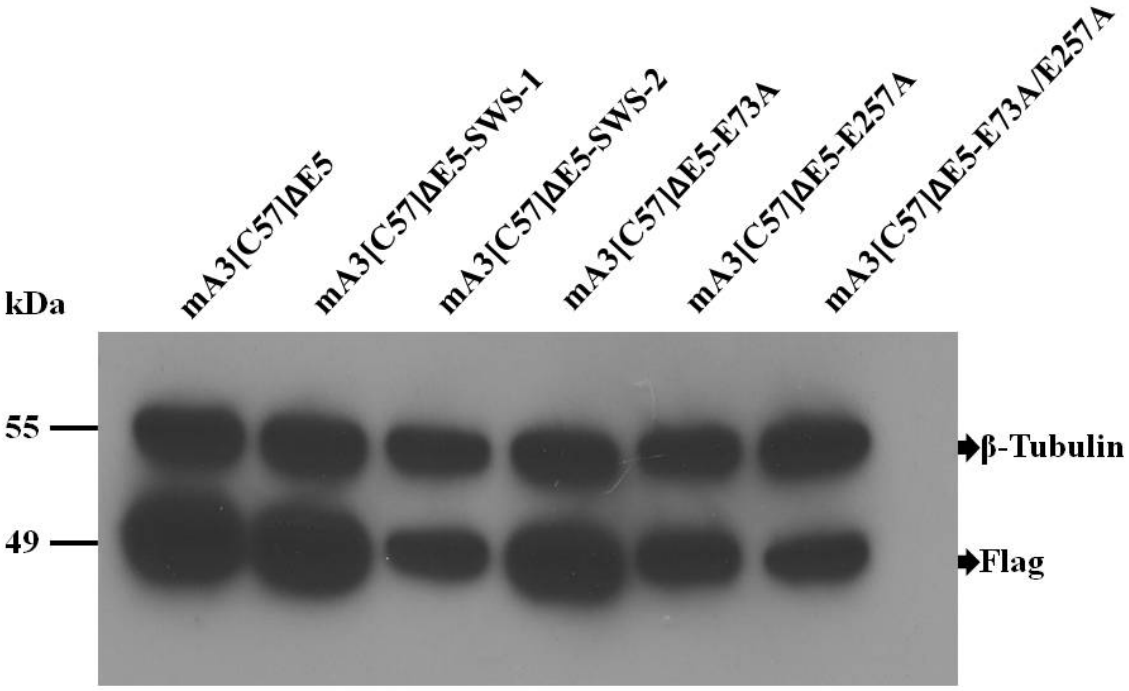

mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]
kDa
55
51
49
β-Tubulin
Flag

Table-2: List of the mA3[C57]ΔE5 point mutants. Point mutants were generated to understand the structural organization of the mA3 protein and the function of the catalytic domains and critical residues within these domains.

mA3 point mutants	Location of the mutation	Potential role of the domain	Expected effect of the mutation
mA3[C57]ΔE5 (wild-type)	-	-	-
mA3[C57]ΔE5-SWS-1	1st Domain	Enzymatic activity	Inhibition of RNA or protein binding
mA3[C57]ΔE5-SWS-2	2nd Domain	RNA binding and encapsidation	Inhibition of RNA or protein binding
mA3[C57]ΔE5-E73A	1st Domain	Enzymatic activity	Inhibition of first catalytic domain
mA3[C57]ΔE5-E257A	2nd Domain	RNA binding and encapsidation	Inhibition of second catalytic domain
mA3[C57]ΔE5-E73A/E257A	1st and 2nd Domain	Enzymatic activity, RNA binding and encapsidation	Inhibition of both catalytic domains

Figure-12: Wild-type mA3[C57]ΔE5 protein with a Flag epitope tag and the point mutants of the wild-type splicing isoform were tested for their expression levels using western immunoblotting. 293T cells were transfected with the Flag-tagged mA3 expression vectors (1 μg) and cell extraction was performed 48 hours post-transfection. The anti-Flag western immunoblotting assay revealed similar expression levels among the wild-type mA3[C57]ΔE5 protein and its mutants.


mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A
kDa
55
49
β-Tubulin
Flag

5.1.3 Retroviral titration assays

The antiretroviral activity of the four mA3 splicing isoforms was tested on six mouse gammaretroviruses and on a HIV-1 pseudovirus that does not express the *Vif* gene (HIVΔvif). All of these viruses encode the eGFP reporter gene to track cell infectivity by flow cytometry analysis. Five of the mouse retroviruses express eGFP from the native LTR promoter, and the HIV-1 and M5P MLV pseudoviruses express eGFP from an internal CMV promoter. Initially, titration analysis was performed on each of these viruses and transducing unit (TU/ml) was calculated to examine their infectivity in tissue culture. A colony represents here a single eGFP-expressing cell.

Six mouse gammaretroviruses (AKV-N, AKV-NB, MoMLV, M5P MLV, Hybrid-1 MLV, and Hybrid-2 MLV) and the HIVΔvif lentivirus were transfected into 293T cells in 6 well plates. One μg of viral plasmid was used for each transfection, and viral particles were collected 48 hours post-transfection to allow the production of adequate amount of viral particles. Supernatants containing the virus from the producer 293T cells were filtered through 0.45μm cartridge filters to avoid potential carryover of the producer cells. The cationic polymer, polybrene, was also added to the samples (8μg/μl) in order to facilitate viral entry in the cell culture, and thus, increase the infectivity levels. Then, tenfold serial dilutions were performed on the filtered supernatants. Mouse NIH 3T3 cells were used as target cells for the mouse gammaretroviruses and human 293T cells were the targets for the HIVΔvif VSV-G-pseudotyped viruses. Target cells were spin-infected for one hour after adding 0.3 ml viral supernatant. Twenty four hours after the spin-infection, target cells were trypsin digested and prepared for flow cytometer analysis. This method was optimized to

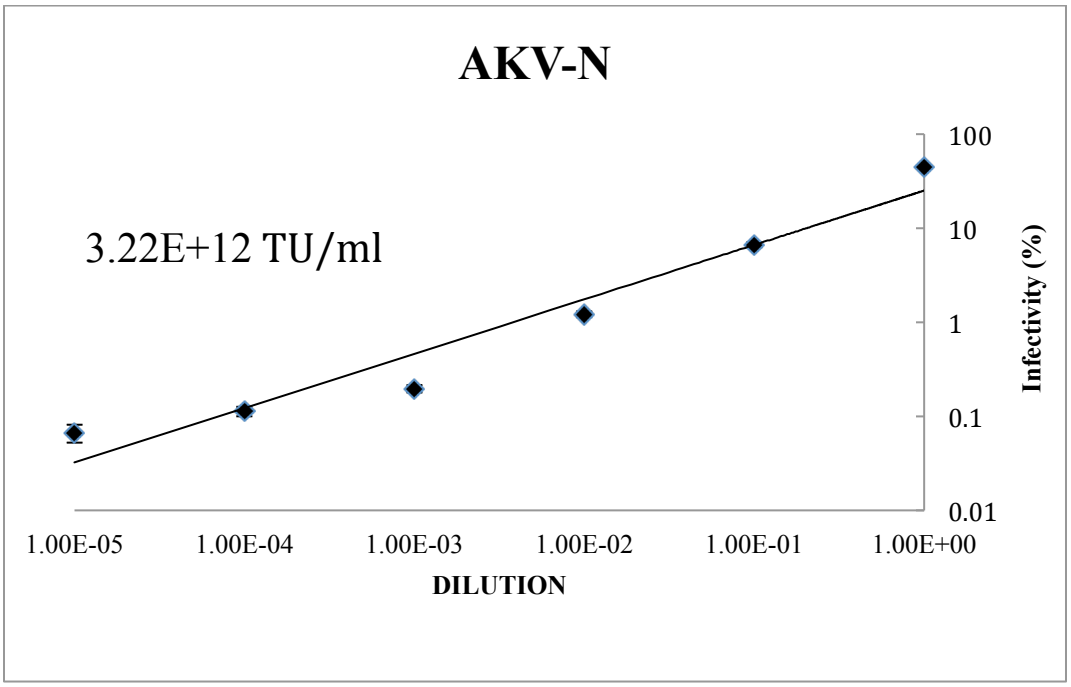
achieve more than 60% viral infection within 24 hours, and to avoid multiple round infections that occur during target cell division.

5.1.3.1 Titration analysis of the AKV-N and AKV-NB gammaretroviruses

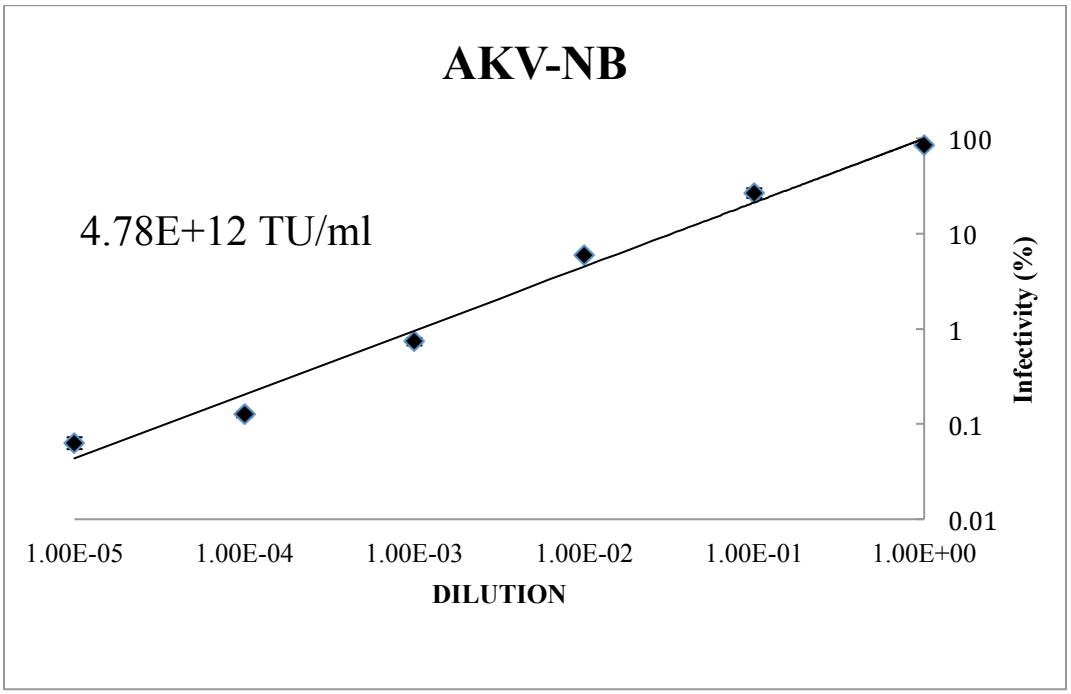
AKV-N is a native endogenous retrovirus of the AKR strain mouse. However, amino acids 1-135 of the gag capsid protein (CA) were replaced with that of MoMLV to generate the AKV-NB virus. This allows the virus to avoid Fv1 restriction in mouse cell lines. Both of these viruses, AKV-N and MoMLV, are replicative mouse gammaretroviruses and it has been identified that one amino acid at position 110 in the CA protein controls the N or B Fv1 tropism of the virus. Alterations in the capsid region allow viruses to overcome Fv1 restriction in a variety of mouse cell lines (1). It should also be mentioned that AKV-NB and MoMLV have more than 80% sequence identity but they behave quite differently in the tissue culture environment, the reasons of which will be addressed in the discussion. Transducing unit analysis for these two viruses, AKV-N and AKV-NB, also indicate that these changes in the CA slightly increased the infectivity of the AKV-NB virus but these two viruses still had the same infection pattern in tissue culture (Figure-13).

Figure-13: Titration Analysis of the AKV-N and AKV-NB mouse gammaretroviruses. Serial tenfold dilutions of the mouse gammaretroviruses were used to infect target NIH 3T3 cells and the infectivity levels were measured 24 hours after the spin-infection. Any cell that contains an eGFP tagged virus will be eGFP positive and Infectivity (%) represents eGFP positive cells. A) The infectivity levels of the AKV-N gammaretrovirus (TU/ml) and the titration analysis for each serial dilution. B) The infectivity levels of the AKV-NB gammaretrovirus (TU/ml) and the titration analysis for each serial dilution.

A)


AKV-N
3.22E+12 TU/ml
100
10
1
0.1
0.01
Infectivity (%)
1.00E-05
1.00E-04
1.00E-03
1.00E-02
1.00E-01
1.00E+00
DILUTION

B)


AKV-NB
4.78E+12 TU/ml
100
10
1
0.1
0.01
Infectivity (%)
1.00E-05
1.00E-04
1.00E-03
1.00E-02
1.00E-01
1.00E+00
DILUTION

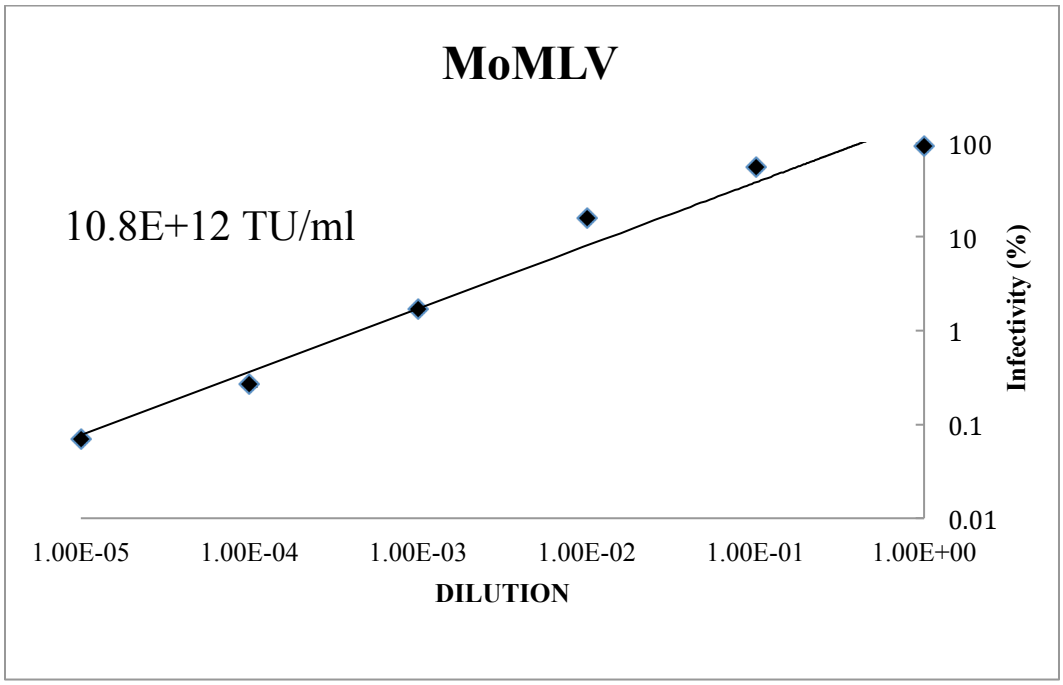
5.1.3.2 Titration analysis of replicative MoMLV and the M5P MLV pseudovirus

The MoMLV and M5P MLV gammaretroviruses are considered to be the same virus but the MoMLV is a replicative virus and M5P MLV is an artificial single-cycle virus that is generated by a multi-plasmid expression system. The VSV-G protein is used in this system to replace the glycoprotein encoded by the *Env* gene of the original virus. The TU analysis of these two viruses was also examined and similar titration levels were detected for both of the viruses (figure-14).

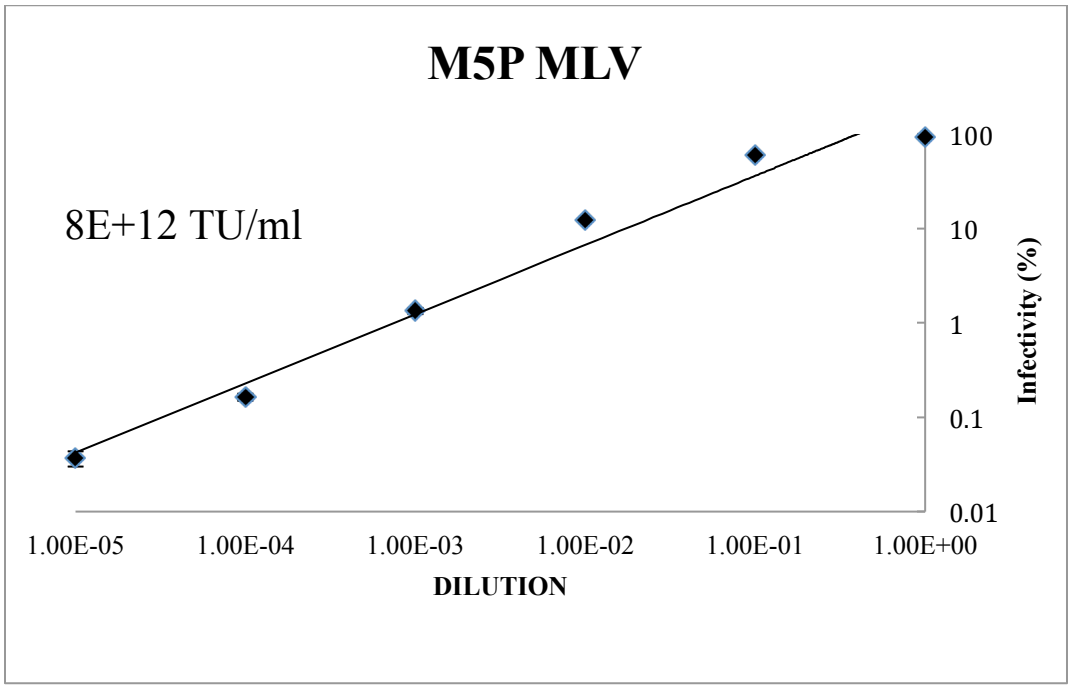
It should also be mentioned that the behaviour of these two viruses in cell culture raised concerns on the fidelity of the multi-plasmid expression system. Although, the titration analysis of the MoMLV and the M5P MLV had similar infectivity levels, subsequent infectivity assays with both of these viruses detected a strong mA3 protein resistance by M5P MLV virus but not with MoMLV. As a result, I concluded that modifications on the viral plasmids should be assessed by more than one experimental approach to reveal the effects of these alterations on the biological activity of the virus. This will be discussed further in the following sections of this chapter.

Figure-14: Titration Analysis of the replicative MoMLV and M5P MLV mouse gammaretroviruses. Serial tenfold dilutions of the mouse gammaretroviruses were used to infect the target 3T3 cells. The infectivity levels were measured 24 hours after the spin-infection. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. A) The infectivity levels of the MoMLV gammaretrovirus (TU/ml) and the titration analysis for each serial dilution. B) The infectivity levels of the M5P MLV gammaretrovirus (TU/ml) and the titration analysis for each serial dilution.

A)


MoMLV
10.8E+12 TU/ml
100
10
1
0.1
0.01
Infectivity (%)
1.00E-05
1.00E-04
1.00E-03
1.00E-02
1.00E-01
1.00E+00
DILUTION

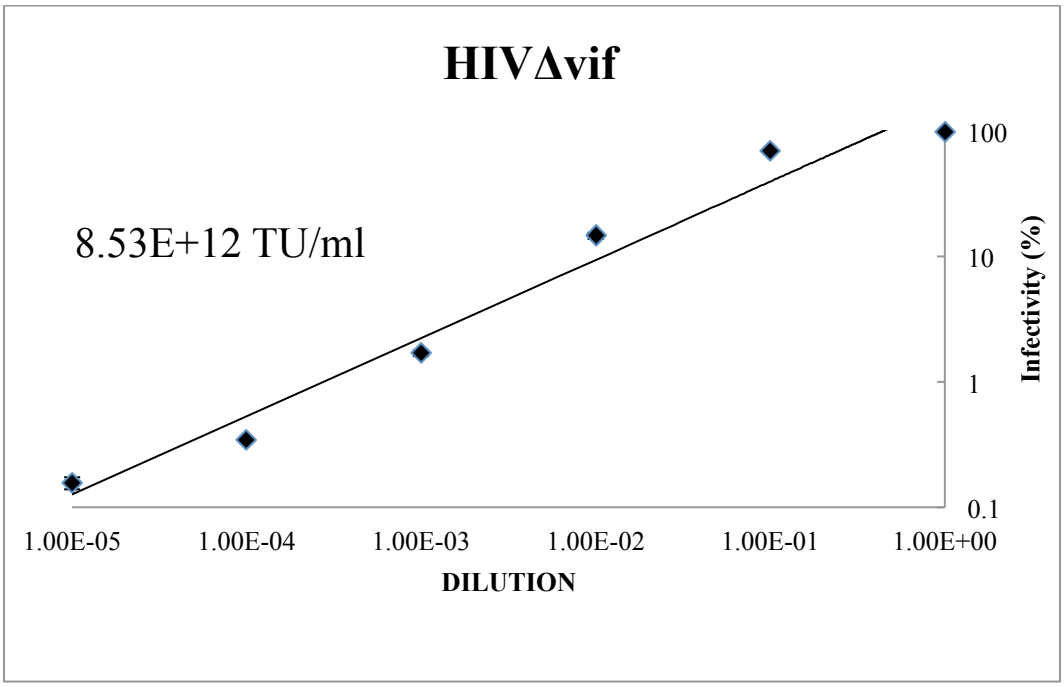
B)


M5P MLV
8E+12 TU/ml
100
10
1
0.1
0.01
Infectivity (%)
1.00E-05
1.00E-04
1.00E-03
1.00E-02
1.00E-01
1.00E+00
DILUTION

5.1.3.3 Titration analysis of the HIVΔvif lentivirus

Understanding the interactions between the mA3 protein and murine gammaretroviruses was one of the main goals of the project. However, replicative murine gammaretroviruses did not allow me to make a proper assessment on various restriction mechanisms of the mA3 protein. The deamination-independent activity of mA3 protein was very strong against these viruses at the early stages of infection, and the absence or low levels of detectable enzymatic activity forced me to test the antiretroviral activity of mA3 protein with a different infection system using a HIVΔvif self-inactivating lentivirus. It was also interesting to test a human retrovirus in my experiments in order to assess the antiretroviral activity of the mA3 protein against a non-murine pathogen. HIV-1 is a lentivirus that targets human $CD4^+$ T cells; however its receptor preference does not allow it to infect murine cells (12). To allow for infection of murine cells, the virus was pseudotyped with VSV-G. The HIV-1 and mA3 protein dynamics can be seen in human cell lines such as 293T cells *in vitro,* allowing for the restriction mechanism of the mA3 protein to act against a human pathogen. The titration analysis for this virus was also performed to assess the infectivity levels in cell culture (Figure-15).

Figure-15: Titration analysis of the VSV-G pseudotyped HIVΔvif lentivirus. Serial tenfold dilution of the HIVΔvif was used to infect the target 293T cells and infectivity levels were measured 24 hours after the spin-infection. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. The infectivity levels of the HIVΔvif lentivirus (TU/ml) and the titration analysis for each serial dilution are shown in the figure.


HIVΔvif
8.53E+12 TU/ml
100
10
1
0.1
Infectivity (%)
1.00E-05
1.00E-04
1.00E-03
1.00E-02
1.00E-01
1.00E+00
DILUTION

5.1.3.4 Conclusions for the protein expression assays and the retroviral titration analysis

The expression levels of the mA3 splicing isoforms were analyzed 48 hours after the transfection when they were at their peak and this time frame was also considered for the infectivity assays (Figure-10A, -10B). Depletion of the mA3 expression vectors was detected after 72 hours and it was concluded that longer incubations might yield different results in the experiments (e.g. less restriction) because of the low expression levels of the mA3 proteins (Figure-10A, -10B). Titration analysis of the retroviruses was necessary to optimize the infectivity of the retroviral pathogens in tissue culture. Expression analysis of the mA3 protein splicing isoforms and results from the retroviral titration assays indicated that 48 hours post-transfection is an ideal time frame for both the production of new viral particles and the optimum protein expression levels. Also, analyzing cells 24 hours post-infection yielded a minimum of 60% viral infection in the target cells (Figure-13, -14, -15). Infection levels in target cells below or above the 60% - 90% range were not considered for our analysis. These conditions were used in future restriction and infectivity assays (Section 5.1.4, 5.2, 5.3, 5.4). In conclusion, this experimental set up allowed me to detect single-cycle retroviral infection and restriction by mA3 during the early stages of infection. Analyzing infected cells 24 hours post-infection allowed me to conduct infectivity assays without saturating the system with high a multiplicity of infection, and as a result, infectivity assays with replicative viruses did not reach the maximum levels in the target cells. In addition, M5P MLV and HIVΔvif pseudoviruses were genetically engineered to generate single-round infections in the target cells and these pseudoviruses were also analyzed 24 hours post-infection.

5.1.4 Comparison of viral restriction among the mA3 splicing isoforms

After establishing an experimental model to assess the antiretroviral activity of mA3 splicing isoforms, co-transfection and infectivity assays were performed to observe the antiretroviral functions of the mA3 splicing isoforms *in vitro.* As it was mentioned in the titration analysis, more than 60% viral infection is necessary to make a significant *in vitro* assessment of the antiretroviral activity of the mA3 proteins. In order to identify the dynamic range of mA3 restriction, 1 μg of the retroviral expression vectors were co-transfected with gradually increasing amounts of mA3 to find the optimum restriction conditions after a single cycle of retroviral replication.

5.1.4.1 Restriction of retroviral pathogens by the mA3[C57]ΔE5 splicing isoform

Viral restriction assays with the mA3[C57]ΔE5 protein and seven different viruses were performed. All of the viruses except M5P MLV had less than 10% infectivity levels when they were co-transfected with 300 ng of mA3[C57]ΔE5 expression vector. I also found a 1:1 mA3-to-retrovirus expression vector ratio was not enough to restrict M5P MLV using mA3[C57]ΔE5 (Figure-16). Similar experimental conditions were used to analyze the antiretroviral capabilities of the other mA3 splicing isoforms and that of the hA3G protein against the AKV-NB virus (Figure-17). These experiments indicated that all the mA3 splicing isoforms, except the low expresser mA3[Balb] were able to cause seven to eight fold restriction with 300 ng of mA3 plasmid DNA in the restriction assays. On the other hand, hA3G was able to restrict the AKV-NB virus when only 100 ng of the hA3G plasmid was used under the same experimental conditions (Figure-17).

Figure-16: Restriction assay with mA3[C57]ΔE5 and different retroviral pathogens. The antiretroviral activity of mA3[C57]ΔE5 was analyzed against six gammaretroviruses and the HIVΔvif pseudovirus. Gradually increasing amounts of the mA3[C57]ΔE5 protein revealed that 300 ng of mA3 expression plasmid is required to inhibit the infectious cycle of the retroviral pathogens 24 hours post infection. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. M5P MLV was not inhibited by the antiretroviral activity of the mA3[C57]ΔE5 protein in the restriction assays.

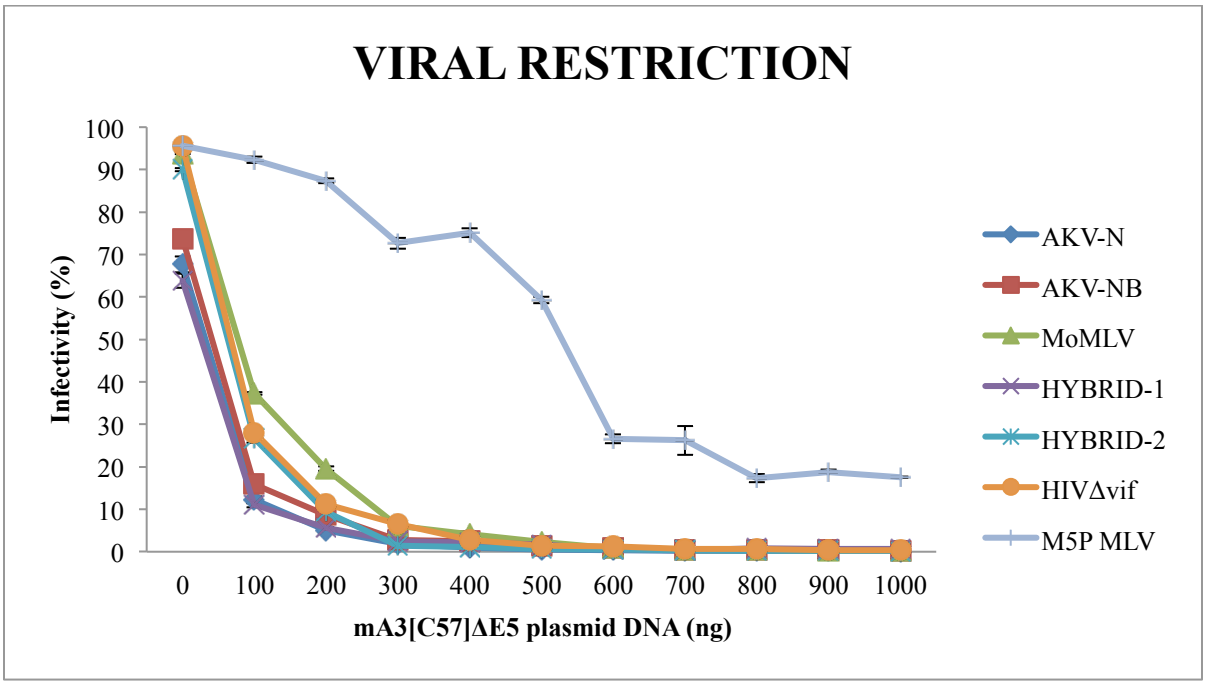
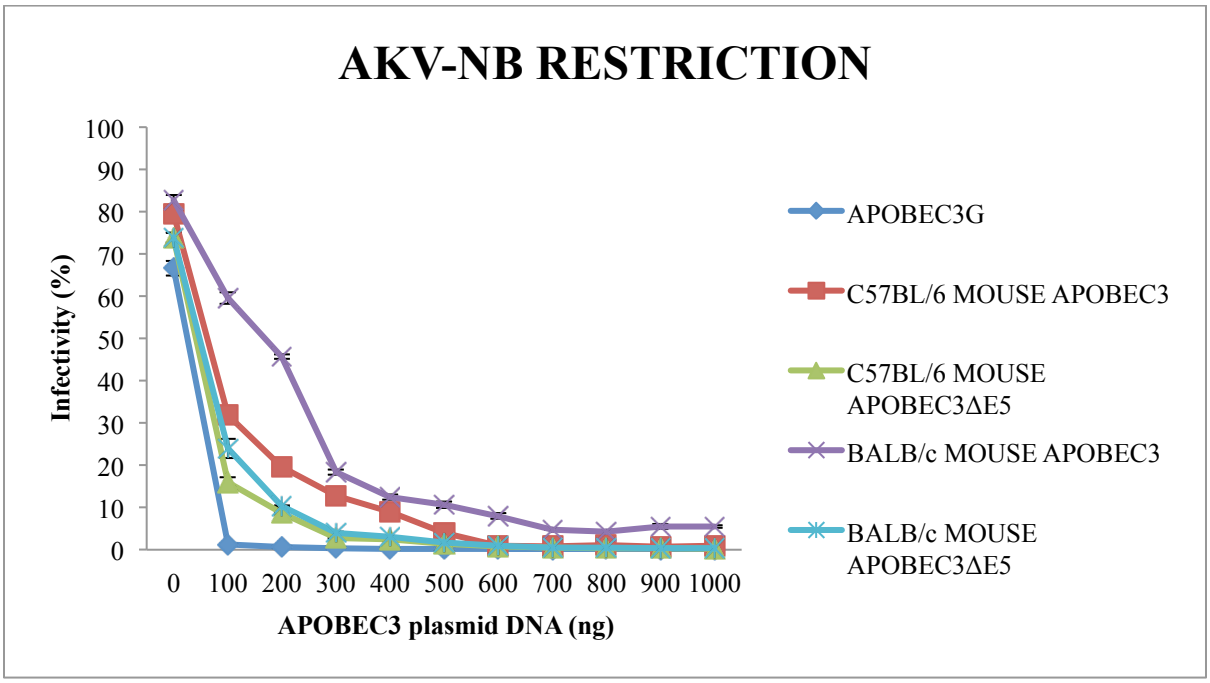

VIRAL RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
0
100
200
300
400
500
600
700
800
900
1000
mA3[C57]ΔE5 plasmid DNA (ng)
AKV-N
AKV-NB
MoMLV
HYBRID-1
HYBRID-2
HIVΔvif
M5P MLV

Figure-17: Restriction of the AKV-NB virus with gradually increasing amounts of the mA3 splicing isoforms and hA3G expression plasmids. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population.The splicing isoforms of the mA3 protein were able to reduce the infectivity levels of the AKV-NB virus down to less than 10% at the 300 ng range and only the low expresser mA3[Balb] splicing isoform had weaker activity against the virus 24 hours post-infection.

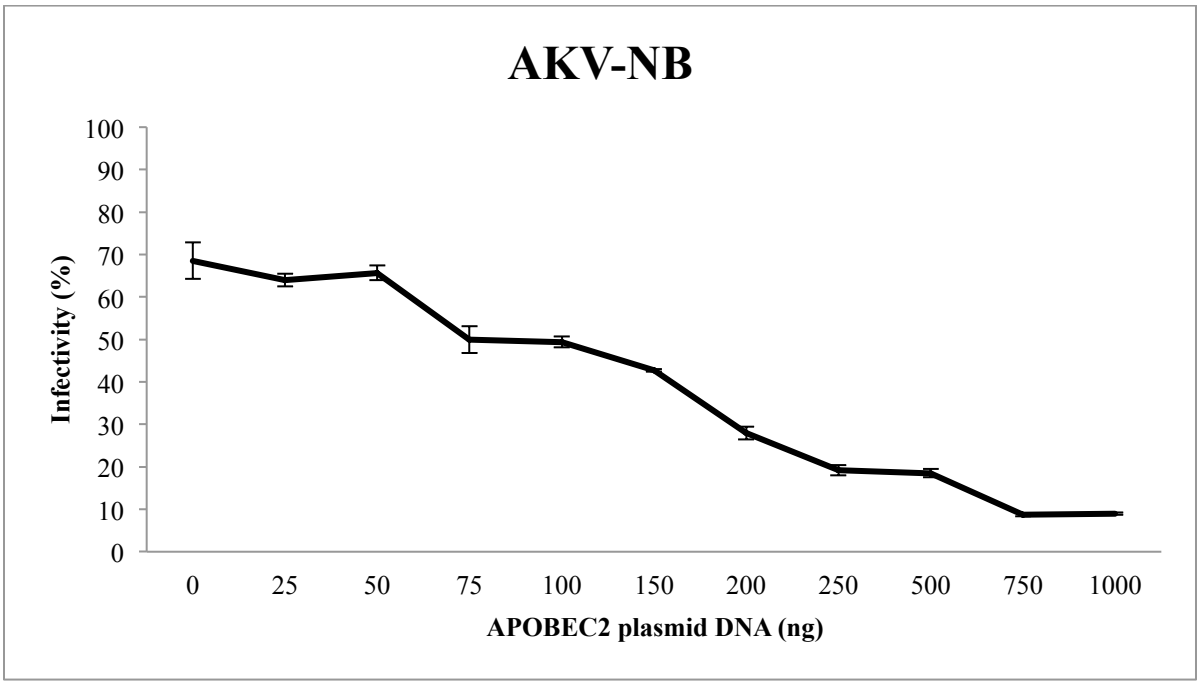

AKV-NB RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
0 100 200 300 400 500 600 700 800 900 1000
APOBEC3 plasmid DNA (ng)
APOBEC3G
C57BL/6 MOUSE APOBEC3
C57BL/6 MOUSE APOBEC3ΔE5
BALB/c MOUSE APOBEC3
BALB/c MOUSE APOBEC3ΔE5

5.1.4.2 Assessment of the A2 protein as a negative control

A2 is also a member of the APOBEC family of proteins but does not have detectable enzymatic or antiretroviral activity against retroviral pathogens (15, 52). This is why I decided to use it as a negative control in our assays. Co-transfection assays with the A2 expression plasmid did not inhibit the infectivity of the retroviral pathogens used in this study. However, expression of the A2 plasmid is extremely high compared to other A3 proteins, as we have found that 25 ng of A2 plasmid yields more protein as per WB analysis than the other A3 proteins at 300 ng dose. The high concentration of the A2 protein in the *in vitro* system sometimes disrupted the infectivity of the retroviral pathogens, and the infection levels of some viruses were negatively affected. The high expression levels of A2 plasmid seemed to reduce the infectivity of AKV-N, AKV-NB and both Hybrid MLV viruses slightly but did not have any negative effects on other viral vectors in the study. However, time course analysis revealed that reduced infectivity levels with A2 can only be observed at the early stages of infection and it is concluded that early effects of A2 expression vector is an experimental artefact that disrupts the *in vitro* infectivity assays for such viruses. The reasons behind the unexpected inhibitory effects of A2 expression vector can be explained by different theories. Firstly, high expression levels of this vector can occupy the cellular ribosomes for protein translation and reduce the amount of competing viral proteins during the translation. Also, high concentrations may disrupt other cellular mechanisms and interfere the viral replication pathways. Finally, different viral plasmids may interact with A2 proteins with an unknown mechanism and slow down the viral kinetics and the replication of less active viral plasmids, such as AKV-N and AKV-NB, can be affected negatively. In order to assess the function of the A2 protein, gradually increasing amounts of the expression

plasmid were tested against the AKV-NB virus in the restriction assays. Unexpectedly, A2 reduced the infectivity levels of the AKV-NB virus at high concentrations (over 100 ng) (Figure-18). This affect was likely a result of the abnormally high expression levels of the A2 protein. As a result, it was decided that 25 ng of the protein could be used as a negative control in the infectivity assays.

Figure-18: Assessment of A2 activity against the AKV-NB virus during in our restriction assays. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. High expression levels of the A2 protein affected the infectivity levels of the virus and the A2 -mediated restriction was observed when the plasmid concentrations increased over 100 ng in the experiments. Only 25 ng of the A2 plasmid was used as a negative control in further infectivity assays.


AKV-NB
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
0
25
50
75
100
150
200
250
500
750
1000
APOBEC2 plasmid DNA (ng)

5.1.4.3 Restriction of pseudoviruses by hA3G

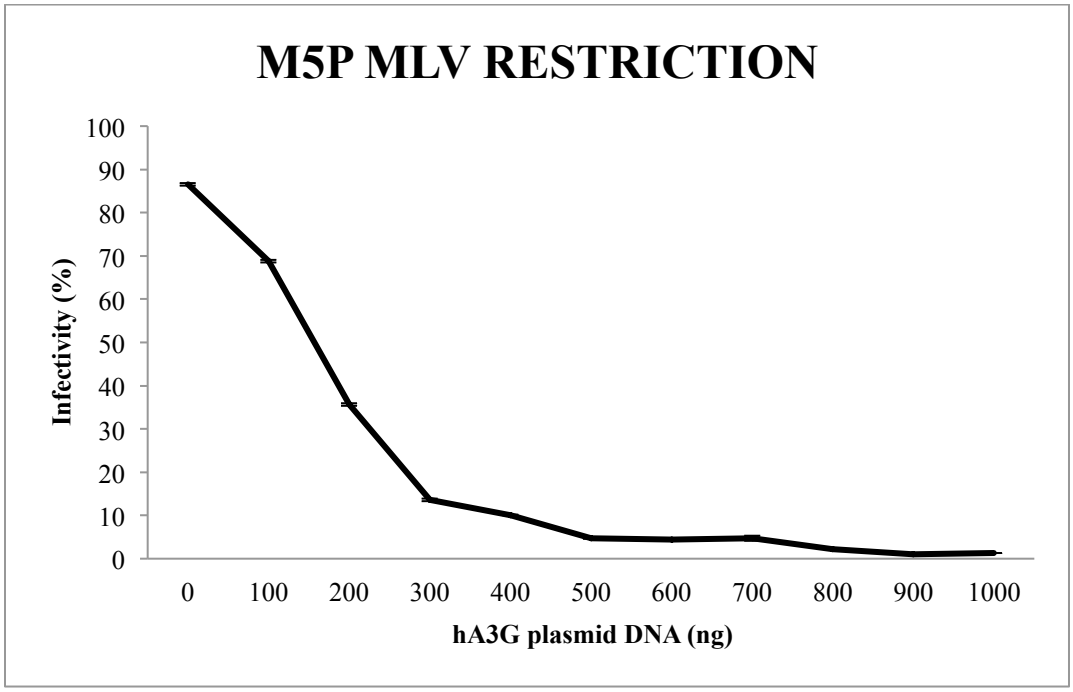
HIVΔvif lentivirus and M5P MLV gammaretrovirus were also tested with increasing amounts of the hA3G plasmid DNA in order to determine the optimum range of the hA3G protein required for restriction of these two viruses *in vitro*. Restriction from the hA3G protein was used as a positive control, since this protein has the most potent antiretroviral activity among the A3 family proteins. Although the infectivity levels of the HIVΔvif virus was reduced to less than 10% with 300 ng of hA3G plasmid DNA (Figure-19B), which was similar to the mA3[C57]ΔE5 restriction on the virus, M5P MLV dropped to this level only when a 1:1 (1µg virus:1µg A3) ratio of hA3G-to-virus plasmid DNA was co-transfected in the experiments (Figure-19A).

5.1.4.4 Conclusions on the restriction assays with the APOBEC family proteins:

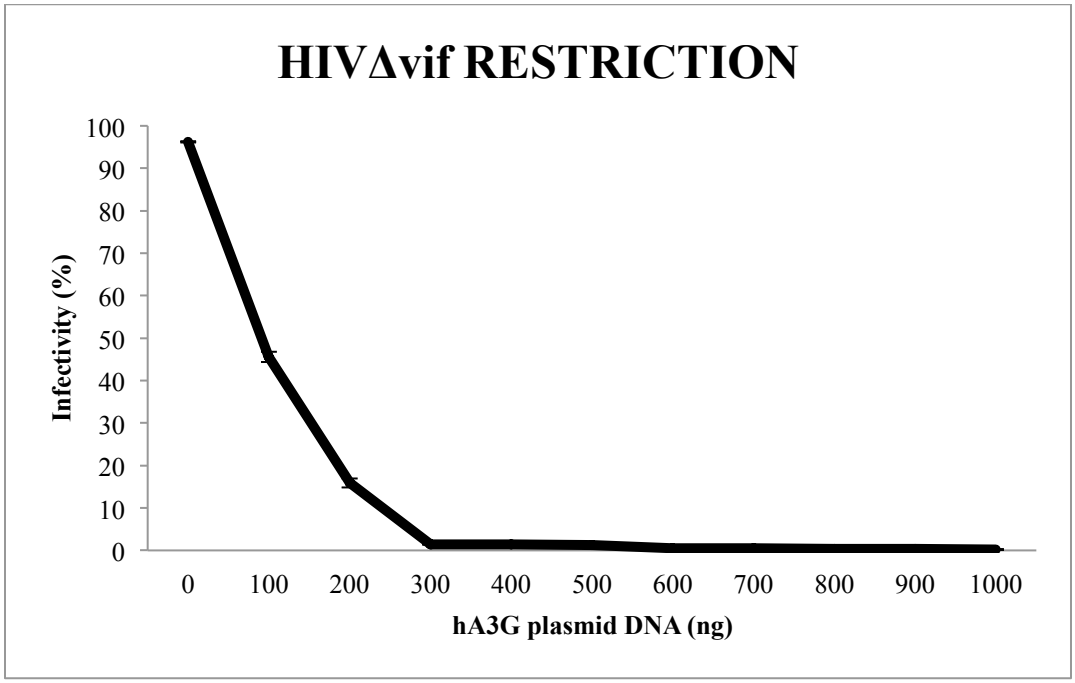
In conclusion, restriction assays with gradually increasing amounts of the A3 plasmid DNA showed that 300 ng plasmid DNA is enough to reduce the infectivity levels of the retroviral pathogens. These are levels also seen in the literature for most of the viruses, except for M5P MLV. The results from these restriction assays were combined with the time course analysis of the mA3 expression levels and viral titration analysis and together, this data set was used in the subsequent experiments as a model to analyze the antiviral activity of mA3 splicing isoforms, and mA3[C57]ΔE5 point mutants.

Figure-19: Assessment of the antiretroviral activity of hA3G against both M5P MLV and HIVΔvif. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. A) M5P MLV was tested against increasing amounts of hA3G. B) HIVΔvif virus was restricted with increasing amounts of hA3G. HIVΔvif was restricted with 300 ng of hA3G, whereas M5P MLV required up to 900 ng for the same complete restriction.

A)


M5P MLV RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
0
100
200
300
400
500
600
700
800
900
1000
hA3G plasmid DNA (ng)

B)


HIVΔvif RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
0
100
200
300
400
500
600
700
800
900
1000
hA3G plasmid DNA (ng)

5.2 Assessment of the antiretroviral activity of mA3 against retroviral pathogens

5.2.1 Antiretroviral activity of the mA3 splicing isoforms and their expression levels in the viral producer cells

In order to investigate the antiretroviral activity of the mA3 splicing isoforms, 300 ng of mA3 was co-transfected with 1 μg of retroviral plasmid DNA. The cells were then incubated for 48 hours to produce the viral particles that are released in the supernatant. These virions have mA3 proteins that were incorporated during the viral assembly. Supernatants from the different transfected cell samples were then collected, filtered to remove the carryover from the producer cells, and used to infect the target cells. These target cells were spin-infected for an hour at 2000 rpm and incubated for another 24 hours before observing the viral infection and/or mA3-mediated retroviral restriction.

In each infectivity assay, three controls have been used to assess the quality and consistency of the experiment. Firstly, 1 μg of retrovirus without A3 was transfected to observe the viral infection lacking restriction activity. Secondly, 25 ng of A2 and 1μg of retrovirus were co-transfected to validate that mA3 proteins are the only source of restriction in the experiments, since A2 does not have any enzymatic or antiretroviral activity. Finally, 100 ng human A3G was co-transfected with 1 μg of retrovirus in the experiments as a positive control to observe the A3-mediated restriction. These three controls have been used throughout the experiments, allowing us to properly compare the antiretroviral activity of the mA3 splicing isoforms while investigating the potency of the mA3-mediated restriction. Moreover, 24 hours after the infection, target cells were trypsinized and prepared for flow cytometer analysis and the antiretroviral activity of the mA3 splicing isoforms as well as control samples were measured

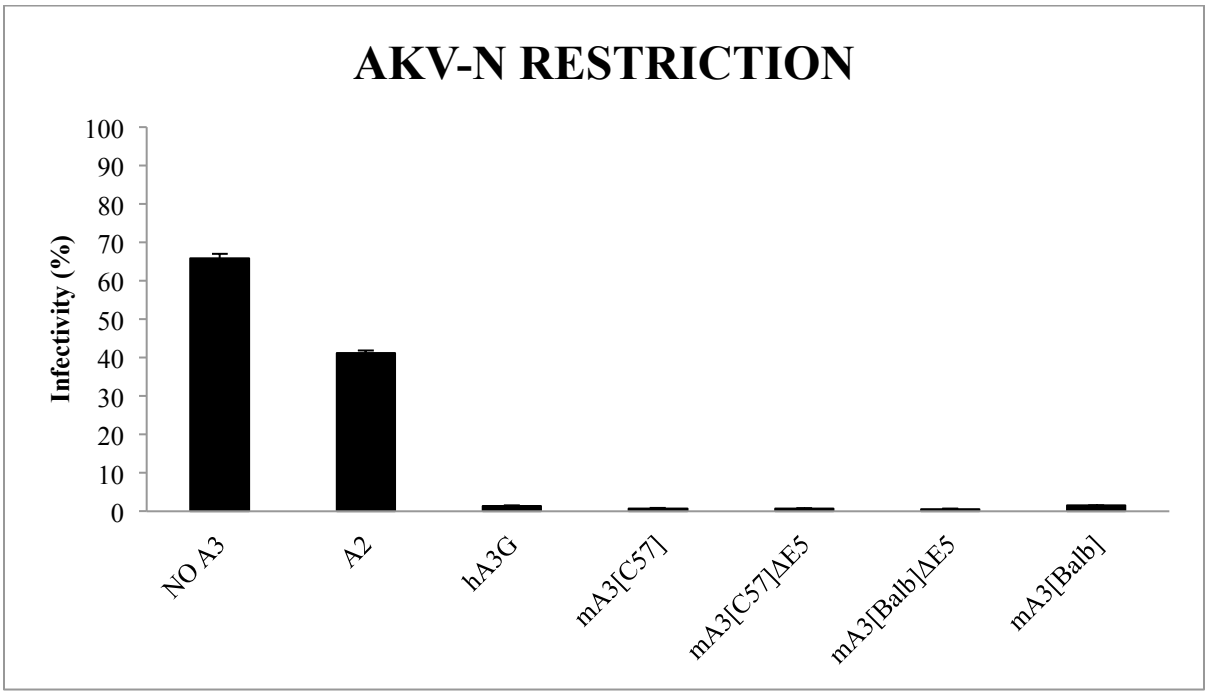
by counting the number of eGFP fluorescent cells in the samples. Twenty thousand cells were counted for each sample in the experiment during the flow cytometry analysis. Producer 293T cell were also harvested and lysed for further protein expression analysis of the Flag-tagged mA3 splicing isoforms, and control samples by western blot.

Regarding the experimental conditions described above, replicative gammaretroviruses (AKV-N, AKV-NB, and MoMLV) and pseudoviruses (M5P MLV and HIVΔvif) were tested against the mA3 splicing isoforms. The results indicated that the murine replicative gammaretroviruses were all inhibited to less than 5% infectivity levels when 300 ng of the mA3 splicing isoforms were co-transfected with these viruses (Figure-20, Figure-21, Figure-22). Even though the low expresser mA3[Balb] protein had little less antiviral effect than the other splicing isoforms, the restriction mediated by this protein was enough to reduce the infectivity levels of virions to less than 10% for all of the replicative gammaretroviruses. As previously mentioned, the M5P MLV is a pseudovirus and unlike the other replicative gammaretroviruses, it contains a CMV promoter instead of an LTR promoter, thereby inducing a faster replication cycle. As a consequence, 300 ng of mA3 expression vector was not enough to inhibit this virus *in vitro* (Figure-23). To resolve this issue, the mA3 expression vector and retrovirus DNA ratio was changed to 1:1, which achieved mA3-mediated restriction (Figure-24) of M5P MLV. Surprisingly, the splicing isoforms of the mA3 protein did not inhibit the infectivity of the HIVΔvif virus equally (Figure-25). The mA3 isoforms lacking exon 5 were potently active against HIVΔvif compared to the full-length isoforms. Western blotting assays confirmed that these different restriction patterns were not related to the mA3 protein levels, as all had similar expression except mA3[Balb] (Figure-25A and -25B). Previous studies also showed the low expression levels of

mA3[Balb] splicing isoform, and indicated that almost five times less mRNA transcripts were produced in the primary cells from BALB/c allele compared to cell from C57BL/6 allele (59). In addition, low expression levels of mA3[Balb] splicing isoforms was considered to be the unstable nature of the protein during the *in vitro* analysis but experiments revealed that even the low expression levels of this splicing isoform was capable of reducing the infectivity of replicative murine retroviruses more than 10 fold.

Figure-20: Antiretroviral activity of the mA3 splicing isoforms against the AKV-N virus *in vitro*. In order to assess the restriction at the early stages of infection, 25 ng A2, 100 ng hA3G and 300 ng mA3 splicing isoforms were co-transfected with 1µg viral plasmid DNA. A) The mA3-mediated restriction of the AKV-N gammaretrovirus was observed after a single-cycle of infection. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. B) Protein expression levels of the control samples and the mA3 splicing isoforms were measured in the 293T producer cells by western immunoblotting. β-Tubulin is a ~55 kDa housekeeping protein, and full-length Flag-mA3 proteins are ~51 kDa and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


AKV-N RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]

B)

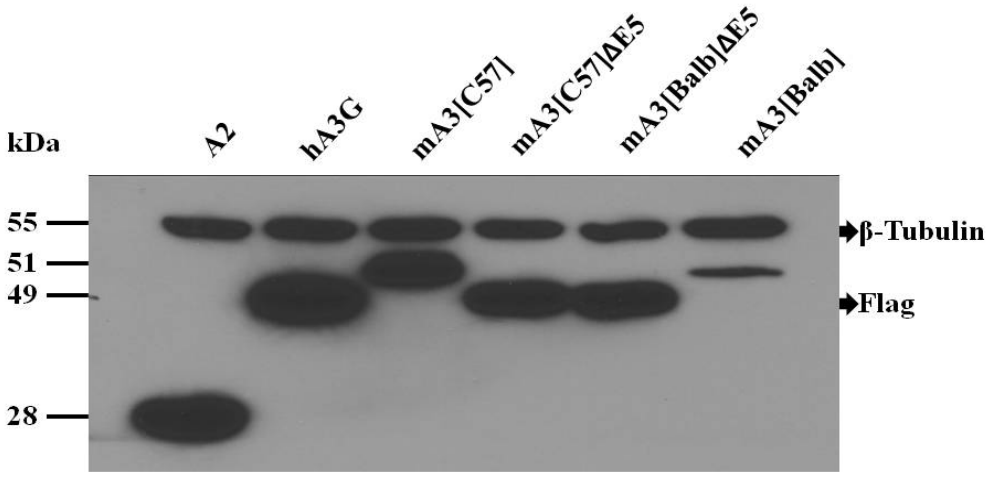
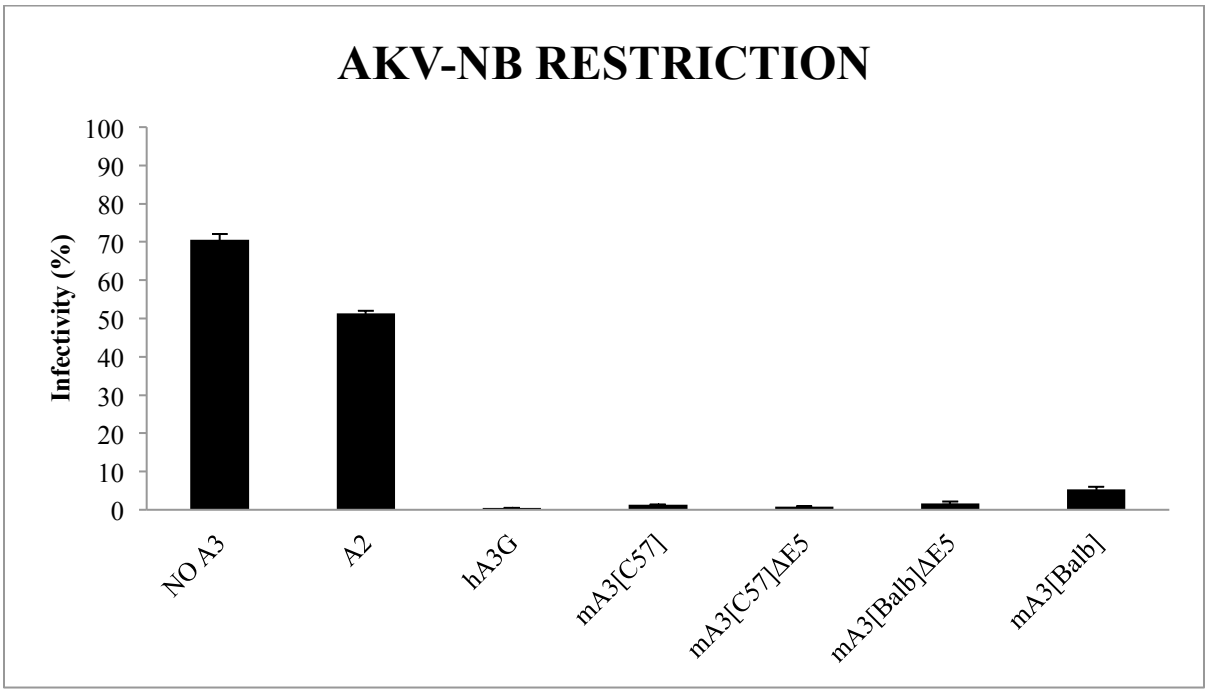

kDa
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]
55
51
49
28
β-Tubulin
Flag

Figure-21: Antiretroviral activity of the mA3 splicing isoforms against the AKV-NB virus *in vitro*. In order to assess the restriction at the early stages of infection, 25 ng A2, 100 ng hA3G and 300 ng mA3 splicing isoforms were co-transfected with 1µg viral plasmid DNA. A) The mA3 -mediated restriction of the AKV-NB gammaretrovirus was observed after a single-cycle of infection. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. B) Protein expression levels of the control samples and the mA3 splicing isoforms were measured in the 293T producer cells by western immunoblotting assay. β-Tubulin is a ~55 kDa housekeeping protein, and full-length Flag-mA3 proteins are ~51 kDa and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


AKV-NB RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]

B)

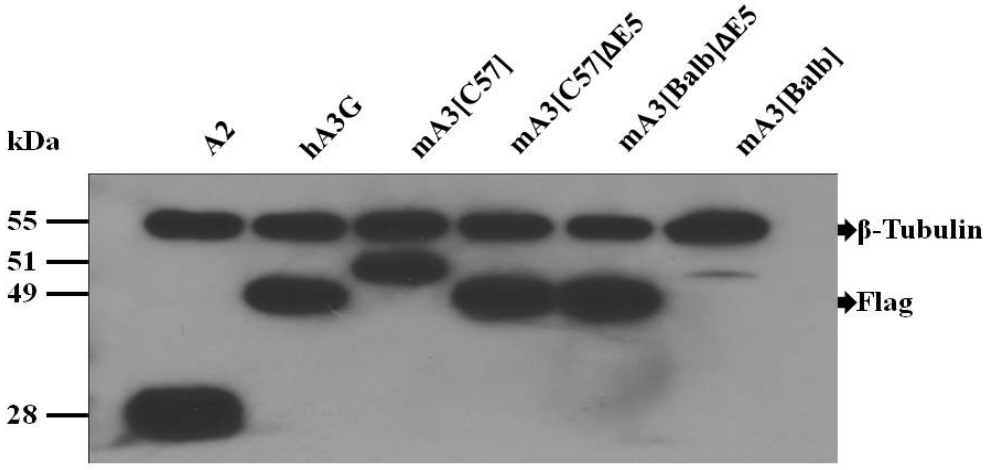
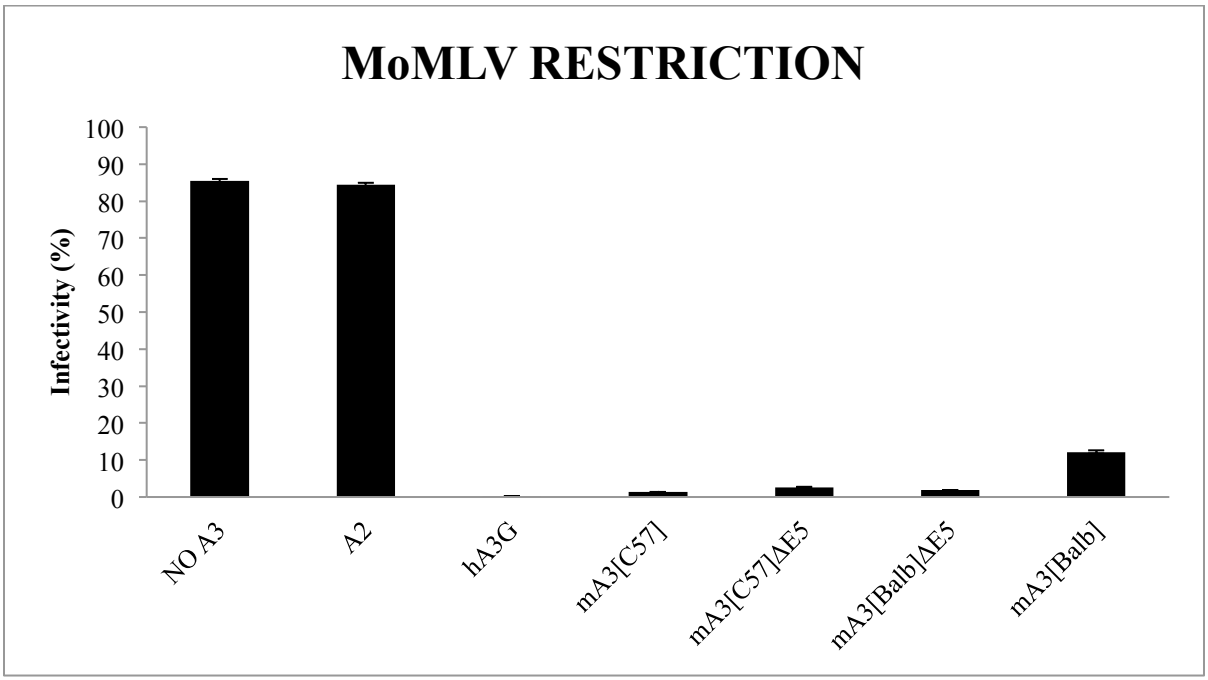

kDa
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]
55
51
49
28
β-Tubulin
Flag

Figure-22: Antiretroviral activity of the mA3 splicing isoforms against the MoMLV *in vitro*. In order to assess the restriction at the early stages of infection, 25 ng A2, 100 ng hA3G and 300 ng mA3 splicing isoforms were co-transfected with 1µg viral plasmid DNA. A) The mA3 -mediated restriction of the MoMLV gammaretrovirus was observed after a single-cycle of infection. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. B) Protein expression levels of the control samples and the mA3 splicing isoforms were measured in the 293T producer cells by western immunoblot. β-Tubulin is a ~55 kDa housekeeping protein, and full-length Flag-mA3 proteins are ~51 kDa and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


MoMLV RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]

B)

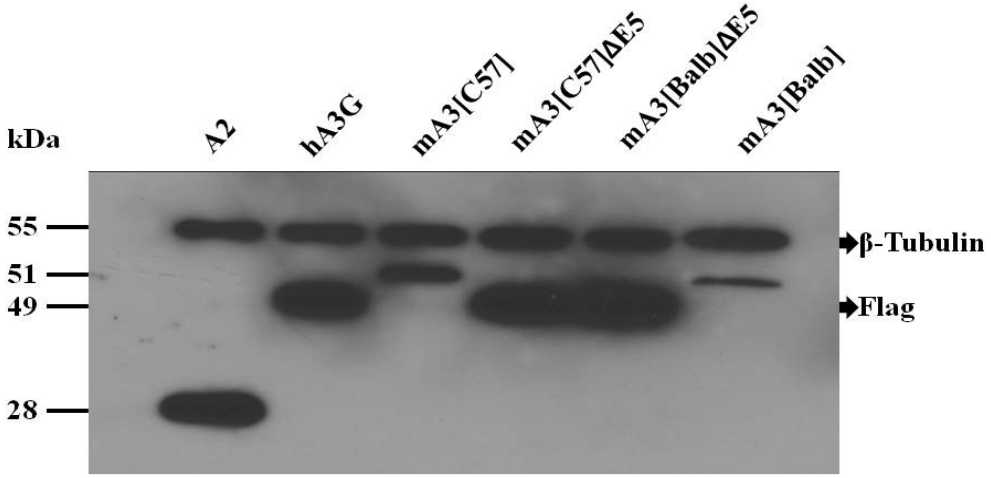
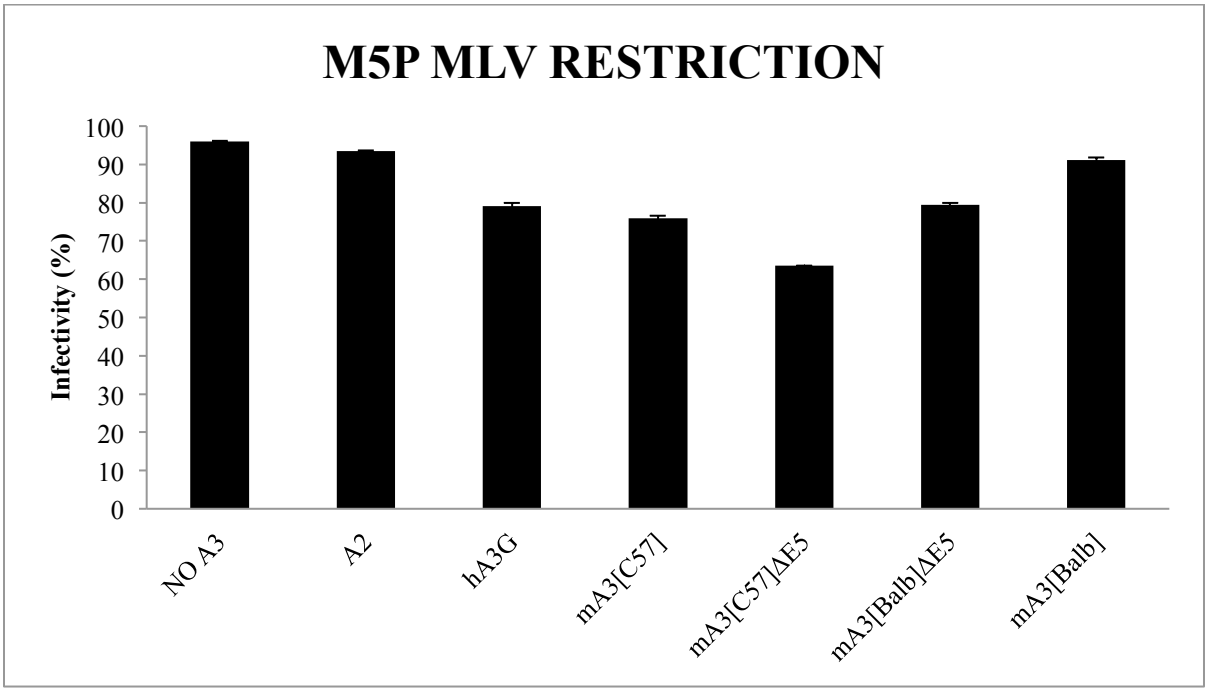

kDa
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]
55
51
49
28
β-Tubulin
Flag

Figure-23: Antiretroviral activity of the mA3 splicing isoforms against M5P MLV *in vitro.* In order to assess the restriction at the early stages of infection, 25 ng A2, 100 ng hA3G and 300 ng mA3 splicing isoforms were co-transfected with 1µg viral plasmid DNA. A) The mA3 -mediated restriction of the M5P MLV pseudovirus was observed after a single-cycle of infection. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. B) Protein expression levels of the control samples and the mA3 splicing isoforms were measured in the 293T producer cells by anti-Flag western immunoblotting assay. β-Tubulin is a ~55 kDa housekeeping protein, and full-length Flag-mA3 proteins are ~51 kDa and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


M5P MLV RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]

B)

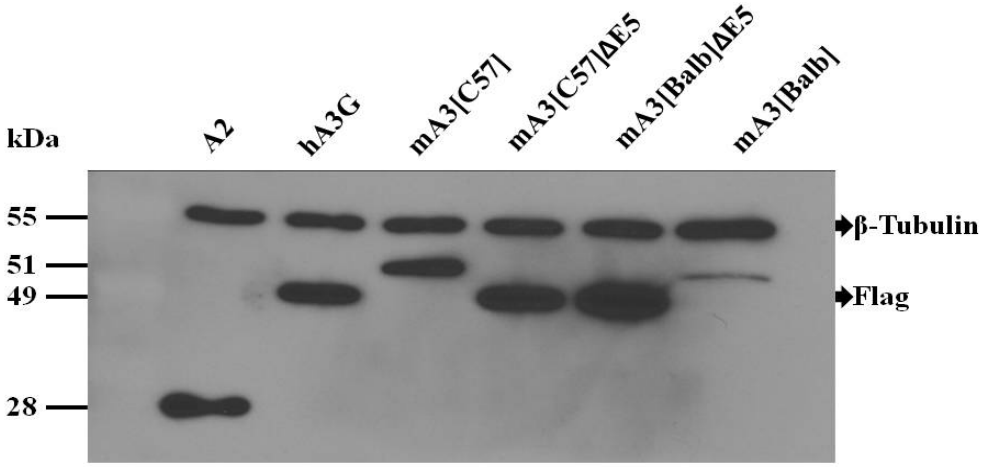

kDa
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]
55
51
49
28
β-Tubulin
Flag

Figure-24: The restriction of the M5P MLV pseudovirus is mediated by higher amounts of mA3 splicing isoforms. The mA3 protein and virus ratio was changed to 1:1 (1µg:1µg) and the infectivity levels of the M5P MLV was reduced to similar levels as the other replicative gammaretroviruses in the *in vitro* experiments.

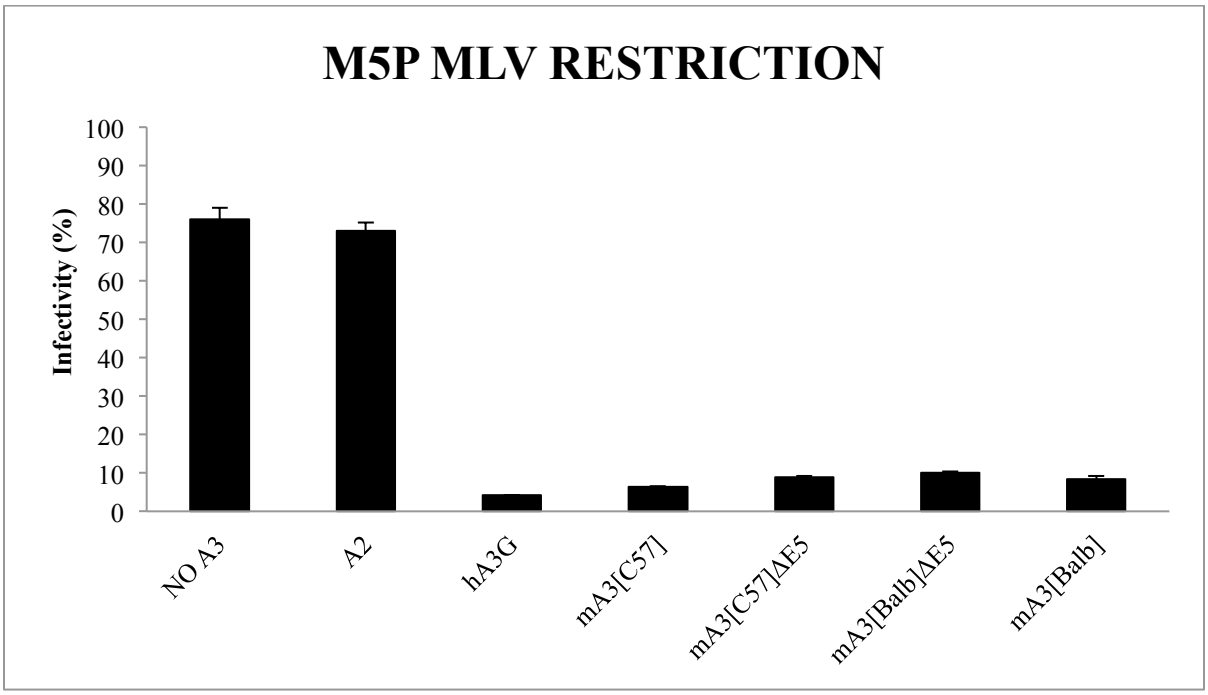
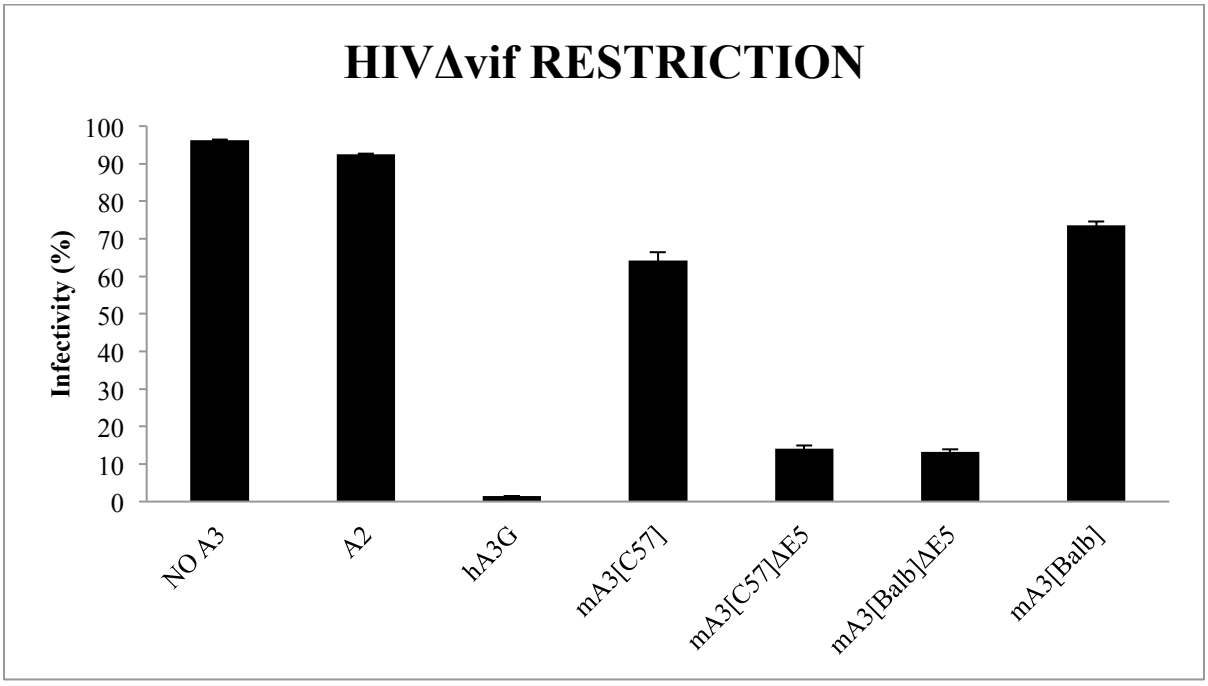
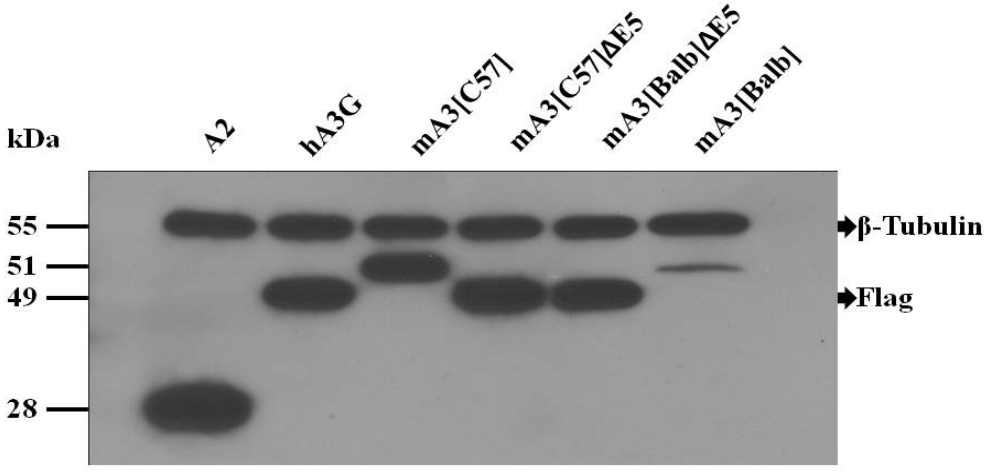

M5P MLV RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]

Figure-25: Antiretroviral activity of the mA3 splicing isoforms against the HIVΔvif lentivirus *in vitro*. In order to assess the restriction at the early stages of infection, 25 ng A2, 300 ng hA3G and 300 ng mA3 splicing isoforms were co-transfected with 1µg viral plasmid DNA. The full-length mA3 protein splicing isoforms were not capable of inhibiting the infectivity of the HIVΔvif lentivirus. A) The mA3 protein-mediated restriction of the HIVΔvif lentivirus was observed after the single-cycle of infection. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. B) Protein expression levels of the control samples and the mA3 splicing isoforms were measured in the 293T producer cells with the anti-Flag western immunoblotting assay. β-Tubulin is a ~55 kDa housekeeping protein, and full-length Flag-mA3 proteins are ~51 kDa and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


HIVΔvif RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]

B)


kDa
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]
55
51
49
28
β-Tubulin
Flag

5.2.2 Functional analysis of the catalytic domains of the mA3[C57]ΔE5 protein

In order to explore the structural organization of mA3 and the functional roles of each of the catalytic domains within the restriction mechanisms, point mutations were generated in the critical residues of the mA3[C57]ΔE5 protein.

Signature motif is indispensable for the enzymatic activity, as proven by various mutational studies (3, 11, 12, 42, 49). Point mutants in the catalytic motif of the N-terminus (mA3[C57]ΔE5-E73A), C-Terminus (mA3[C57]ΔE5-E257A), and both catalytic domains (mA3[C57]ΔE5-E73A/E257A), were used to investigate the role of each domain on deamination-mediated restriction. Since the SWS nucleic acid binding motifs found in both the N- and C- termini of the hA3G protein are thought to play a crucial antiretroviral role, these regions were mutated within the mA3[C57]ΔE5 protein to explore their functions. Investigations of the SWS nucleic acid binding motifs within the mA3[C57]ΔE5 protein revealed that the SWS-2 point mutant in the C- terminus of the protein may contribute slightly to the antiretroviral activity of the protein, but overall, these mutants did not have the same drastic effect on the restrictive capabilities of the mA3[C57]ΔE5 protein compared to the hA3G protein. The antiretroviral activity of the mA3[C57]ΔE5 protein point mutants were measured against the AKV-N, AKV-NB, MoMLV, M5P MLV, and HIVΔvif viruses in order to observe the restriction trends of these mutants against a variety of retroviral pathogens.

5.2.2.1 The antiretroviral activity of the mA3[C57]ΔE5 protein point mutants against the replicative gammaretroviruses

Experiments on the mA3[C57]ΔE5 protein and its point mutants indicated that the antiviral activity of these mutant proteins still remained against the replicative gammaretroviruses (AKV-N, AKV-NB, MoMLV). Although the point mutants of the first domain (mA3[C57]ΔE5-E73A), the second domain (mA3[C57]ΔE5-E257A) and both domains (mA3[C57]ΔE5-E73A/E257A) of the mA3[C57]ΔE5 protein were generated to potentially impair the enzymatic activity of that domain, these mutant proteins remained active against the replicative gammaretroviruses, albeit with fairly less antiviral activity than the wild-type mA3[C57]ΔE5 protein (Figure-26, Figure-27, Figure-28). It was concluded that the antiretroviral activity of the mA3[C57]ΔE5 protein against the replicative gammaretroviruses is a result of the deamination-independent restriction pathway.

Figure-26: Determination of the functional roles of the catalytic domains and the critical residues of the mA3 protein *in vitro*. In order to assess the restriction during the early stages of infection, 25 ng A2, 100 ng hA3G and 300 ng mA3 splicing isoforms were co-transfected with 1μg viral plasmid DNA. A) The antiretroviral activity of the wild-type mA3[C57]ΔE5 protein and its point mutants were tested against the AKV-N gammaretrovirus to investigate the enzymatic activity and the functional roles of the catalytic domains. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. B) Protein expression levels of the control samples and the mA3[C57]ΔE5 protein point mutants were measured in the 293T producer cells with the anti-Flag western immunoblotting assay. β-Tubulin is a ~55 kDa housekeeping protein, and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


AKV-N RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A
hA3G-E69A/E259A

B)

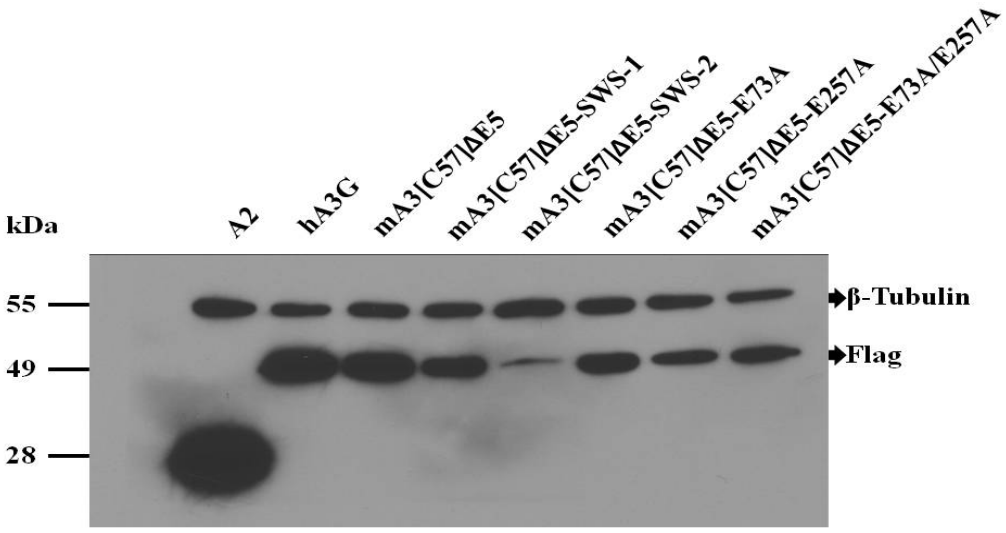
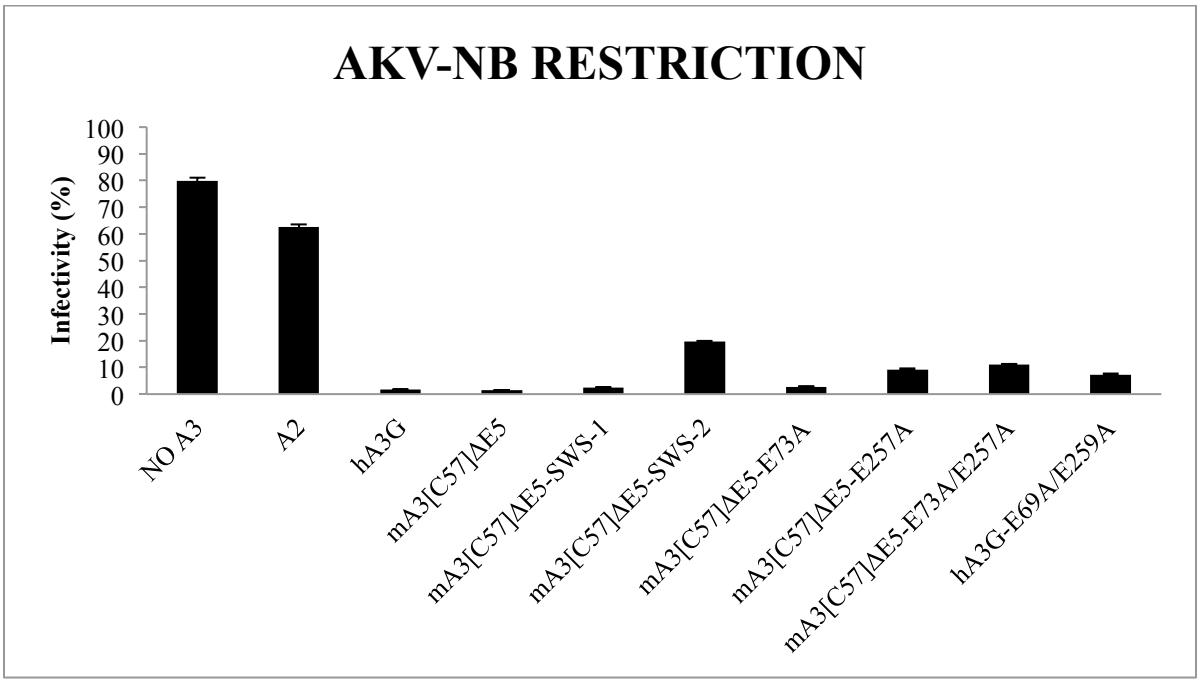

kDa
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A
55
49
28
β-Tubulin
Flag

Figure-27: Determination of the functional roles of the catalytic domains and the critical residues of the mA3 protein *in vitro*. In order to assess the restriction at the early stages of infection, 25 ng A2, 100 ng hA3G and 300 ng mA3 splicing isoforms were co-transfected with 1μg viral plasmid DNA. A) The antiretroviral activity of the wild-type mA3[C57]ΔE5 protein and its point mutants were tested against the AKV-NB gammaretrovirus to investigate the enzymatic activity and the functional roles of the catalytic domains. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. B) Protein expression levels of the control samples and the mA3[C57]ΔE5 protein point mutants were measured in the 293T producer cells with the anti-Flag western immunoblotting assay. β-Tubulin is a ~55 kDa housekeeping protein, and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


AKV-NB RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A
hA3G-E69A/E259A

B)

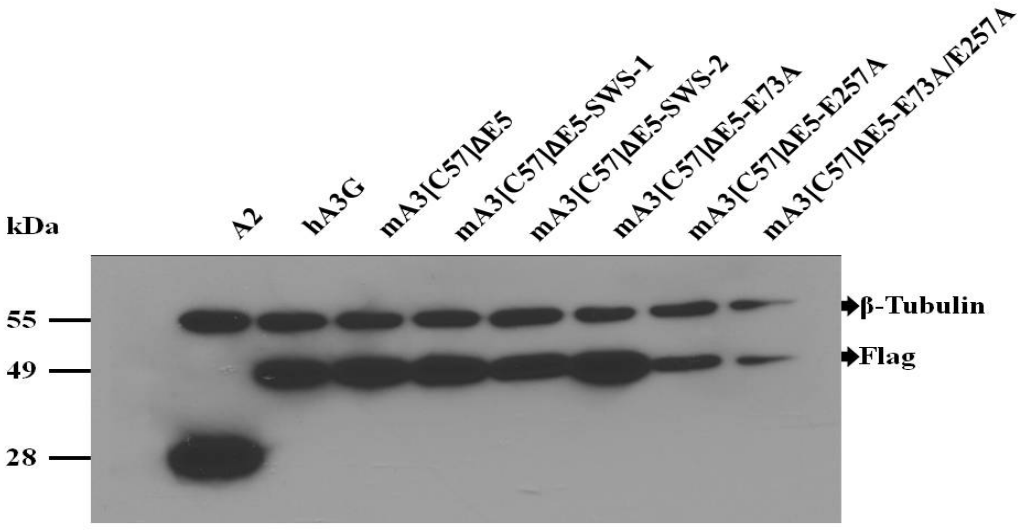
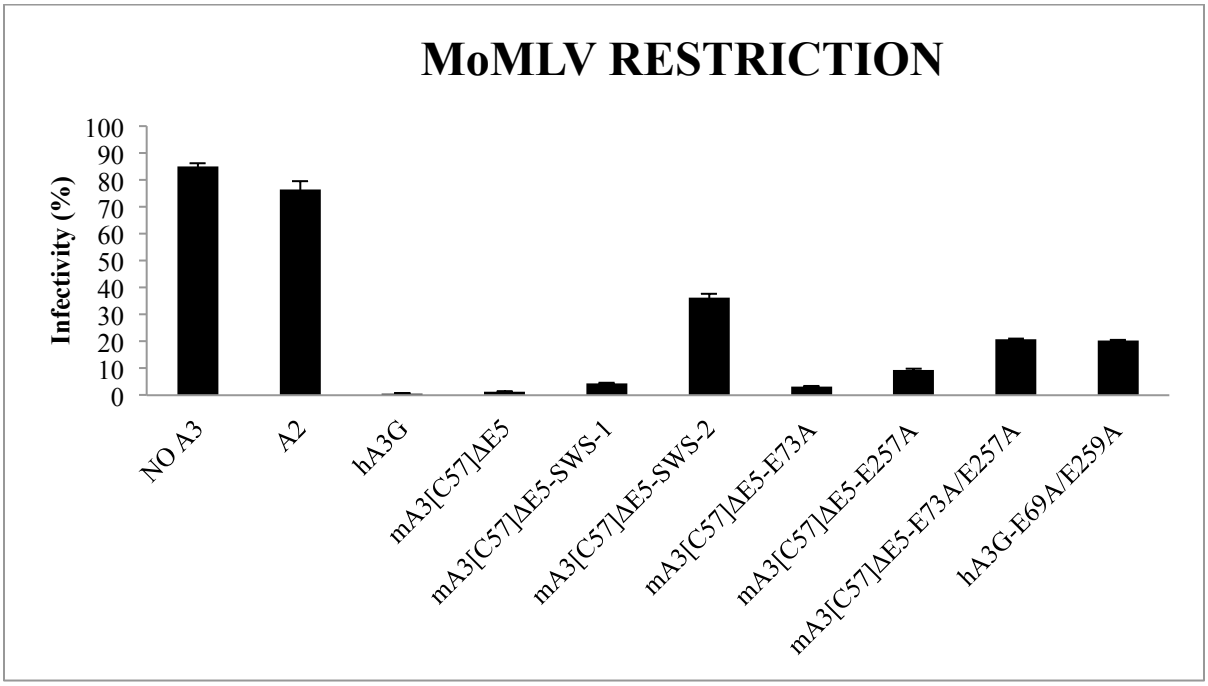
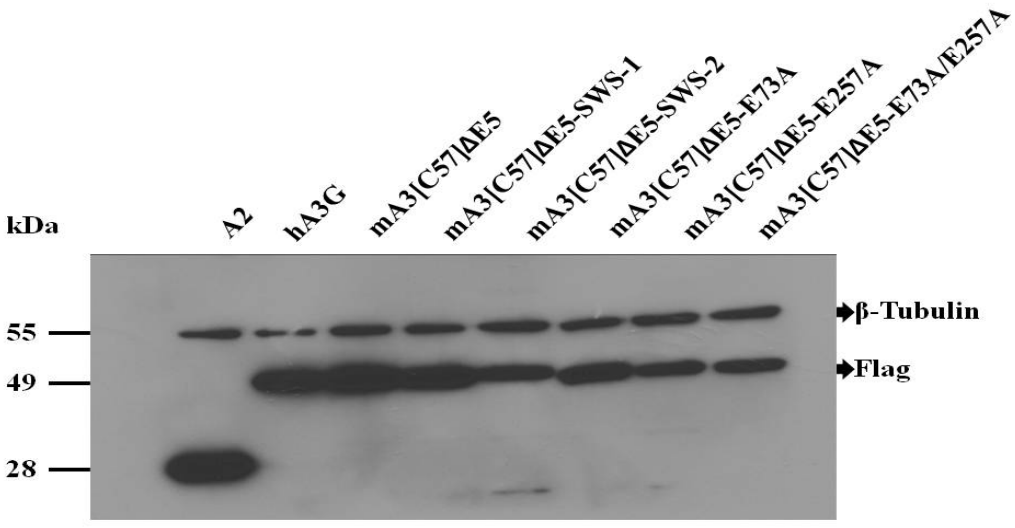

kDa
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A
55
49
28
β-Tubulin
Flag

Figure-28: Determination of the functional roles of the catalytic domains and the critical residues of the mA3 protein *in vitro*. In order to assess the restriction at the early stages of infection, 25 ng A2, 100 ng hA3G and 300 ng mA3 splicing isoforms were co-transfected with 1μg viral plasmid DNA. A) The antiretroviral activity of the wild-type mA3[C57]ΔE5 protein and its point mutants were tested against the MoMLV replicative gammaretrovirus to investigate the enzymatic activity and the functional roles of the catalytic domains. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. B) Protein expression levels of the control samples and the mA3[C57]ΔE5 protein point mutants were measured in the 293T producer cells with the anti-Flag western immunoblotting assay. β-Tubulin is a ~55 kDa housekeeping protein, and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


MoMLV RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A
hA3G-E69A/E259A

B)


kDa
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A
55
49
28
β-Tubulin
Flag

5.2.2.2 The antiretroviral activity of the mA3[C57]ΔE5 point mutants against M5P MLV and HIVΔvif pseudoviruses

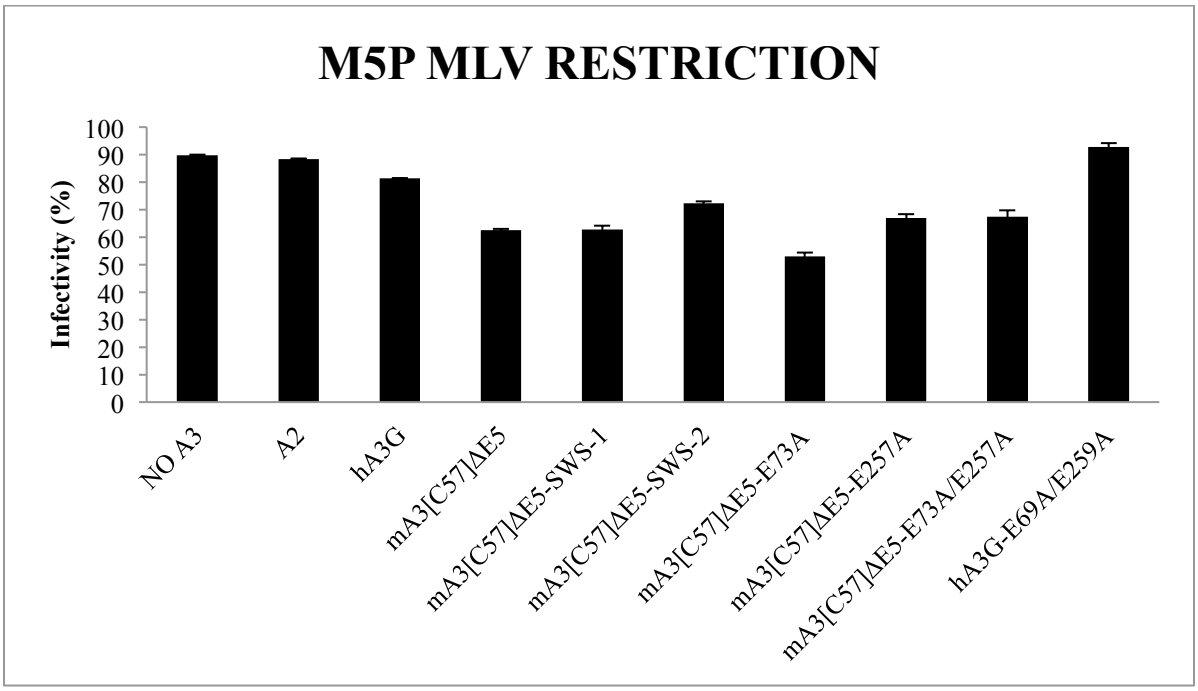
The low doses (300 ng) of mA3 against M5P MLV again failed to restrict the virus (Figure-29). The mA3: virus plasmid ratio was set to 1:1 (1µg:1µg) again and the experiment was repeated with higher quantities of mA3, causing infectivity levels to drastically drop (Figure-30). Also, the restriction levels of M5P MLV from the second experiment were similar to what was observed with the MoMLV and the mA3[C57]ΔE5 protein point mutants in the infectivity assays (Figure-28). However, mA3[C57]ΔE5-SWS-2 mutant was still less active against M5P MLV pseudovirus and it was hypothesized that different structural and genomic organization of the pseudoviruses may provide resistance to these viruses, and also, genetic modifications on A3 proteins may cause unexpected effects on some point mutants. Overall, the activity of mA3[C57]ΔE5-SWS-2 mutant was reduced against M5P MLV pseudovirus but it was still able to decrease the infectivity levels more than 30% in the assays.

The results from the HIVΔvif infectivity assay also showed that the point mutants of the first and second catalytic domains remained partially active against the virus, but did not completely inhibit the infectious cycle of the virus to the minimum levels seen with h3G. When both domains of the mA3[C57]ΔE5 protein were mutated there was still detectable amounts of antiretroviral activity but this activity was not enough to reduce the infectivity levels to less than 40 % as it was shown with replicative murine gammaretroviruses It was concluded that this activity is not a deamination-mediated restriction mechanism of the mA3 protein (Figure-31). Also, the A2 control revealed that *in vitro* expression levels of the proteins alone does not disrupt the infectivity of this virus. In general, the antiretroviral activity of the point mutants was considerably weaker against the HIVΔvif virus, indicating

that the deamination-independent restriction pathway may have a weaker activity against this retroviral pathogen or the virus may have increased infectivity *in vitro*.

Figure-29: Determination of the functional roles of the catalytic domains and the critical residues of the mA3 protein *in vitro*. In order to assess the restriction at the early stages of infection, 25 ng A2, 100 ng hA3G and 300 ng mA3 splicing isoforms were co-transfected with 1μg viral plasmid DNA. A) The antiretroviral activity of the wild-type mA3[C57]ΔE5 protein and its point mutants were tested against the M5P MLV pseudovirus to investigate the enzymatic activity and the functional roles of the catalytic domains. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. B) Protein expression levels of the control samples and the mA3[C57]ΔE5 protein point mutants were measured in the 293T producer cells with the anti-Flag western immunoblotting assay. β-Tubulin is a ~55 kDa housekeeping protein, and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


M5P MLV RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A
hA3G-E69A/E259A

B)

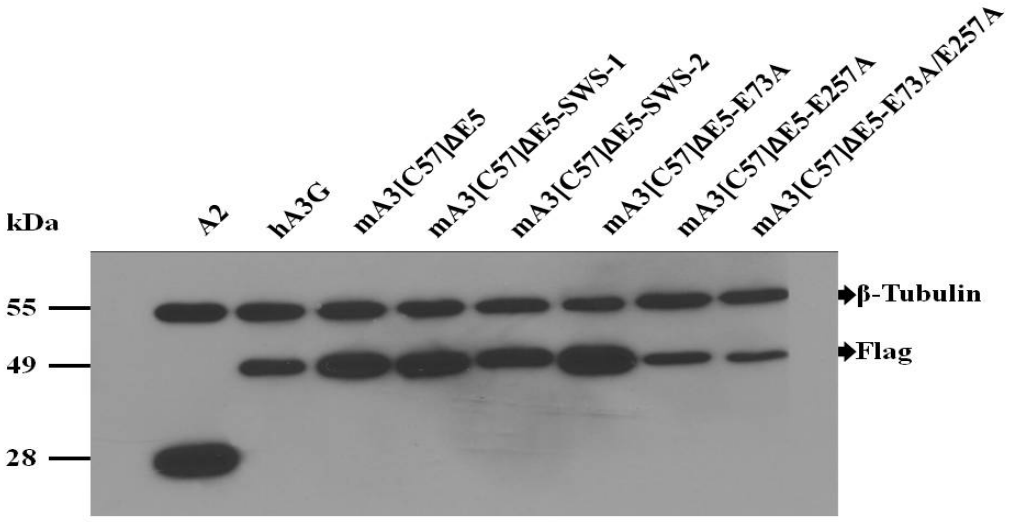

kDa
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A
55
49
28
β-Tubulin
Flag

Figure-30: Restriction of the M5P MLV pseudovirus with increased quantities of the mA3[C57]ΔE5 point mutants. The mA3-to-virus ratio was set to 1:1 again in order to decrease the infectivity levels of the M5P MLV to the adequate levels and to be able to observe the antiviral function of the mA3[C57]ΔE5 protein point mutants in the experiments.

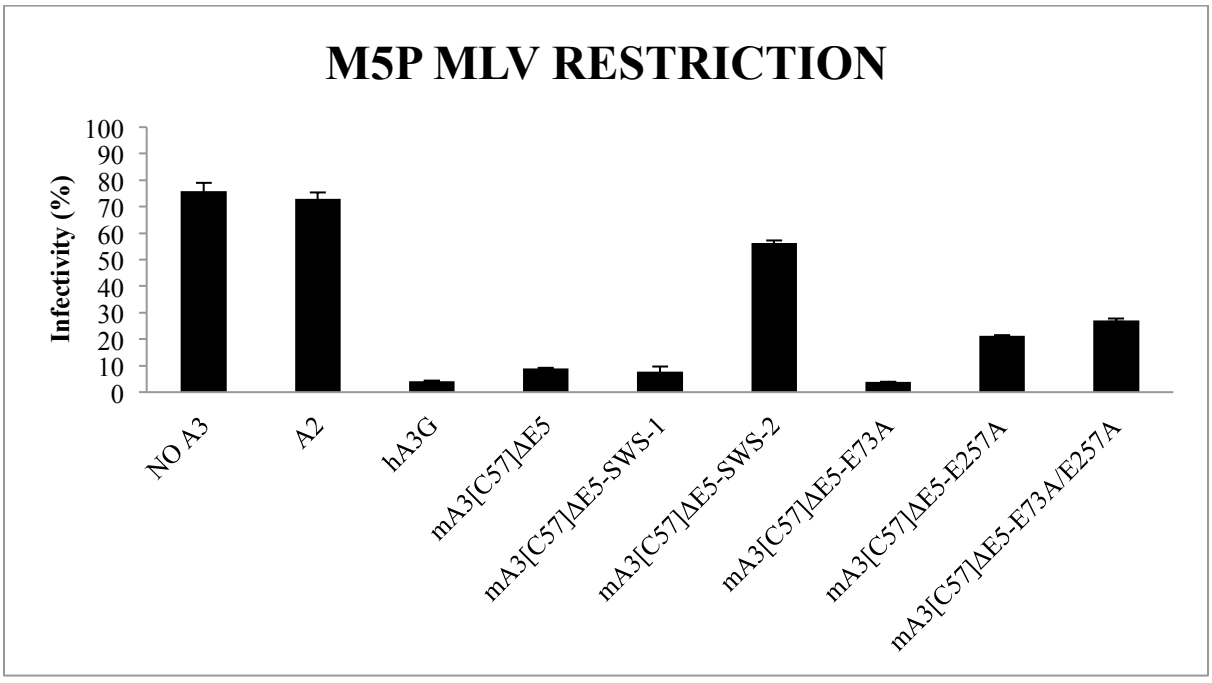
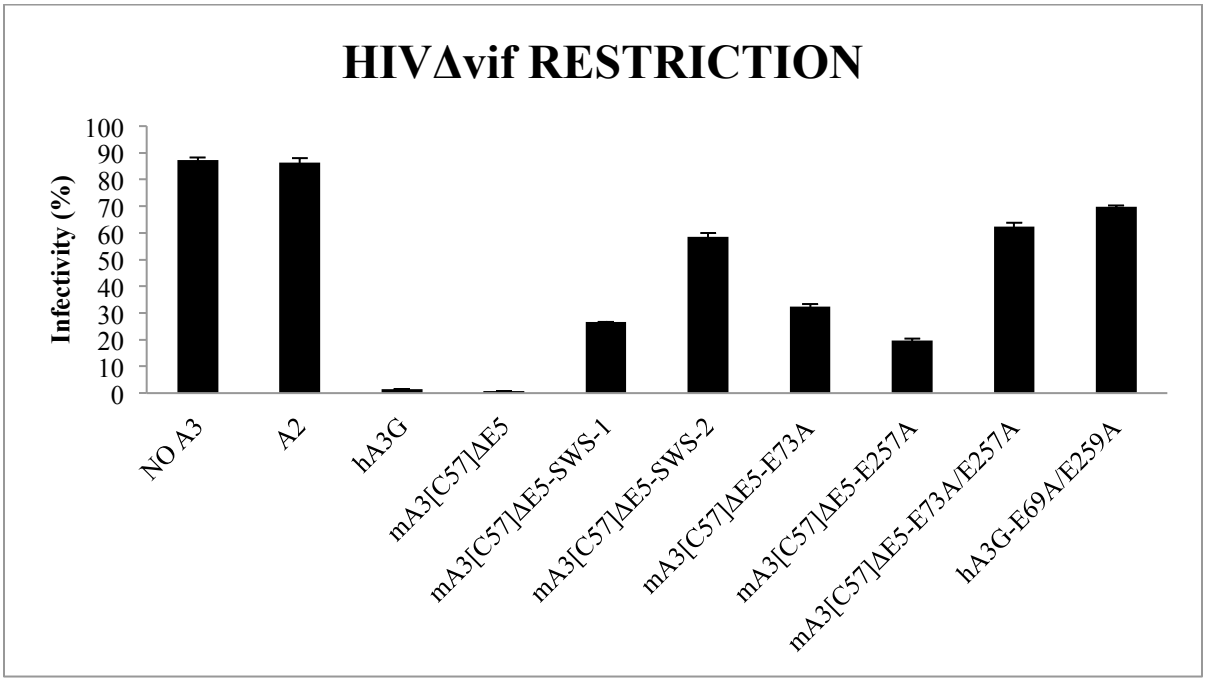
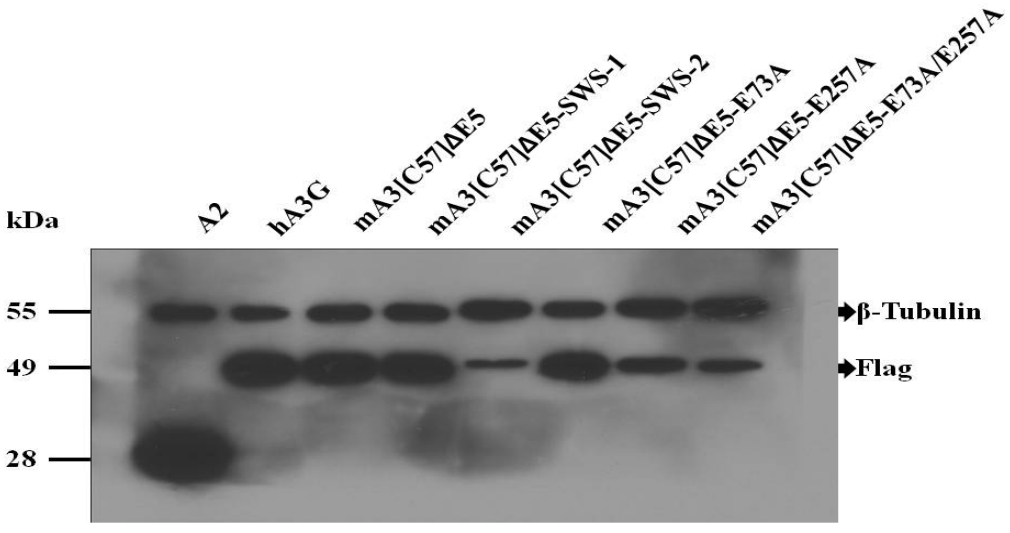

M5P MLV RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A

Figure-31: Determination of the functional roles of the catalytic domains and the critical residues of the mA3 protein *in vitro*. In order to assess the restriction during the early stages of infection, 25 ng A2, 300 ng hA3G and 300 ng mA3 splicing isoforms were co-transfected with 1μg viral plasmid DNA. A) The antiretroviral activity of the wild-type mA3[C57]ΔE5 protein and its point mutants were tested against the HIVΔvif lentivirus to investigate the enzymatic activity and the functional roles of the catalytic domains. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. B) Protein expression levels of the control samples and the mA3[C57]ΔE5 protein point mutants were measured in the 293T producer cells with the anti-Flag western immunoblotting assay. β-Tubulin is a ~55 kDa housekeeping protein, and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


HIVΔvif RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A
hA3G-E69A/E259A

B)

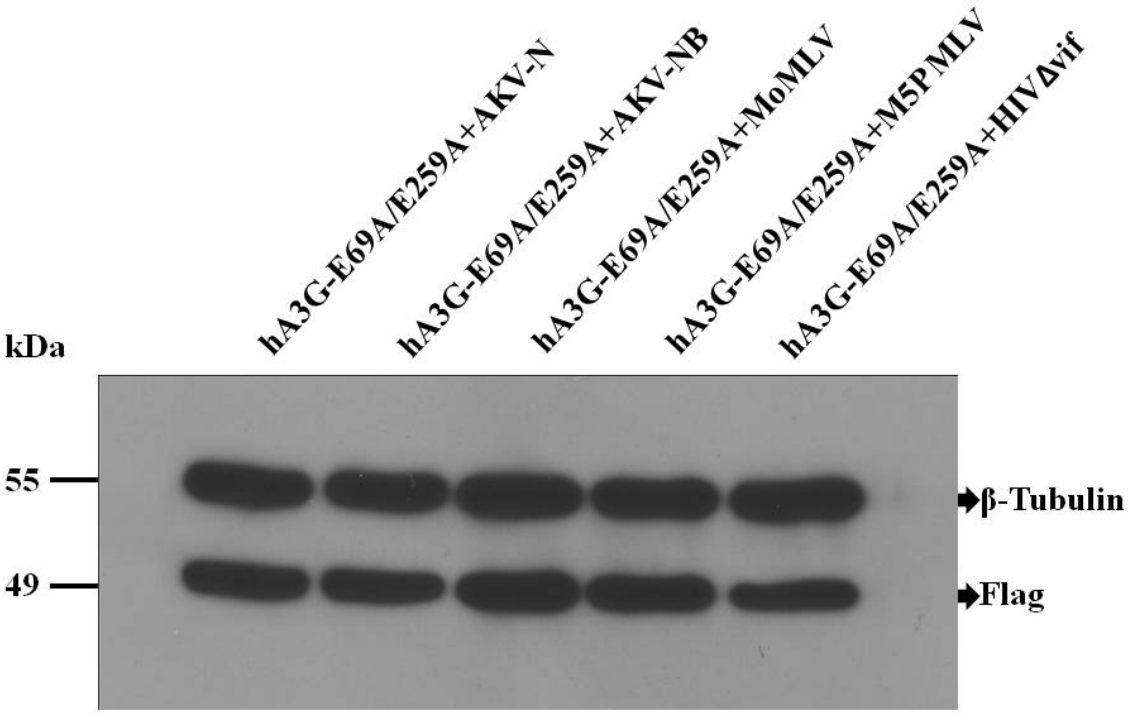

kDa
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-SWS-1
mA3[C57]ΔE5-SWS-2
mA3[C57]ΔE5-E73A
mA3[C57]ΔE5-E257A
mA3[C57]ΔE5-E73A/E257A
55
49
28
β-Tubulin
Flag

5.2.2.3 Conclusions on the functional roles of the mA3 protein catalytic domains

Regarding the structural organization and the functional roles of the catalytic domains of the mA3 protein, I have shown that both domains are involved in the restriction mechanisms of the protein and the antiretroviral activity of the mA3 catalytic domains is not limited to the deamination-dependent restriction mechanisms. A strong deamination-independent restriction during the early stages of infection is able to drastically reduce the infectivity levels of retroviral pathogens *in vitro*. Furthermore, the point mutants of the mA3[C57]ΔE5 protein also exhibited a deamination-independent pathway as observed in the experiments. Due to this potent antiviral activity, it was hard to distinguish which domain was needed for enzymatic activity of the mA3 protein. Thus, it was decided that further sequencing analysis of the proviral genomes in the target cells is required to assess the enzymatic activity of the protein. This would allow me to increase the sensitivity of our investigation regarding the source of antiviral activity.

Moreover, after the initial observations of the deamination-independent restriction pathway of the mA3 protein, catalytically dead hA3G protein (hA3G-E69A/E259A) were also generated and the antiviral activity of this protein was compared with its murine counterpart in the experiments (Figure-26, -27, -28, -29, -30, -31). The results indicated that the hA3-E69A/E259A protein also had strong deamination-independent activity against gammaretroviruses, and this activity was reduced against the HIVΔvif virus. The expression level of the protein was also measured against the retroviral pathogens in the experiments (Figure-32).

Figure-32: Protein expression levels of hA3G-E69A/E259A in the restriction assays with murine gammaretroviruses and HIVΔvif. In order to assess the restriction during the early stages of infection (24 hours), 300 ng hA3G-E69A/E259A plasmid DNA was co-transfected with 1μg viral plasmid DNA. Producer cells were harvested after the collection of the supernatant for the infectivity assays and expression levels of hA3G-E69A/E259A was measured by anti-Flag western-immunoblotting assay. β-Tubulin is a ~55 kDa housekeeping protein, and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.


hA3G-E69A/E259A+AKV-N
hA3G-E69A/E259A+AKV-NB
hA3G-E69A/E259A+MoMLV
hA3G-E69A/E259A+M5PMLV
hA3G-E69A/E259A+HIVΔvif
kDa
55
49
β-Tubulin
Flag

5.2.3 HMM and LMM complex formation of the mA3 splicing isoforms and the mA3[C57]ΔE5 protein point mutants

A3 proteins are able to form HMM and LMM complexes within the cytoplasm depending on the current situation of the cell. When A3 proteins assemble into HMM complexes, their enzymatic activity is impaired and they are no longer available to deaminate the DNA sequences in mononucleotide context (12, 47, 70). In contrast, their catalytic activity to edit DNA sequences is not affected when they are in LMM complexes (12, 47, 70). As a result, any modifications in the A3 protein structural organization may have a potential effect on A3 complex formation and antiviral activity of the protein. I thought it was interesting to investigate whether these complexes were altered by using the mA3 protein splicing isoforms and the point mutants of the mA3[C57]ΔE5 protein.

The mA3 expression vectors were transfected into 293T cells and incubated for 48 hours to reach the optimum protein expression levels. Two days after the transfection, 293T cells were harvested and prepared for velocity sedimentation. A 40% to 5% sucrose gradient was prepared for each sample prior to cell extraction, and cell lysates were overlaid onto the gradient solution. Samples were then spun for 6 hours at 41K RPM and each sucrose gradient was separated and prepared for anti-Flag western immunoblotting assays to determine the molecular mass changes.

The results of these sedimentation assays showed that all of the mA3 protein splicing isoforms (Figure-33) and the mA3[C57]ΔE5 protein point mutants (Figure-34) assembled into HMM complexes as expected. In conclusion, I show that A3 proteins form HMM complexes in the rapidly dividing 293T cell line and it is clear that splicing isoforms and point mutants do not disrupt the formation of these complexes in these cells.

Figure-33: Molecular mass complex formation of the mA3 splicing isoforms *in vitro*. The same amount of mA3 splicing isoform expression vectors (1 μg) were transfected into 293T cells and incubated 48 hours, after which they were harvested and applied to a velocity sedimentation assay. Each layer was then analyzed with western immunoblot. The mA3 splicing isoforms shifted to the bottom of the sucrose gradient (40%) in the velocity sedimentation assays showing they are found in HMM complexes. Samples were analyzed with the anti-Flag western immunoblotting assays in order to determine the changes in the molecular mass formation. β-Tubulin is a ~55 kDa housekeeping protein, and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

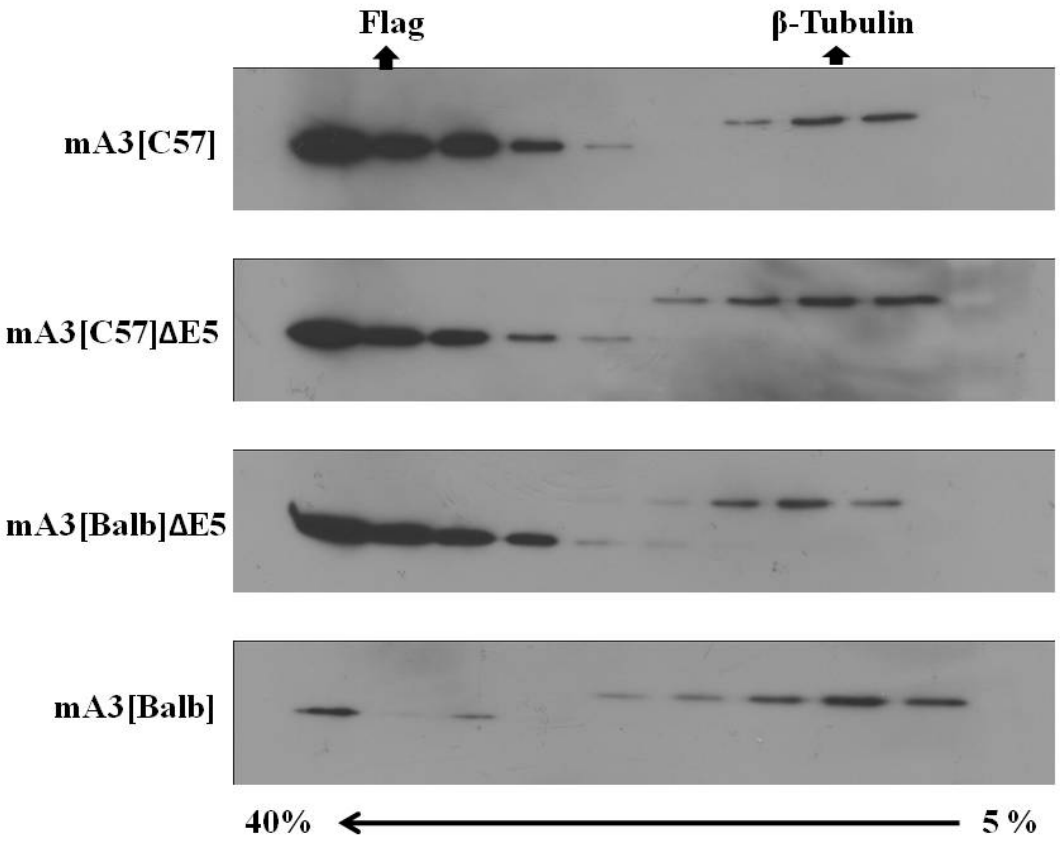
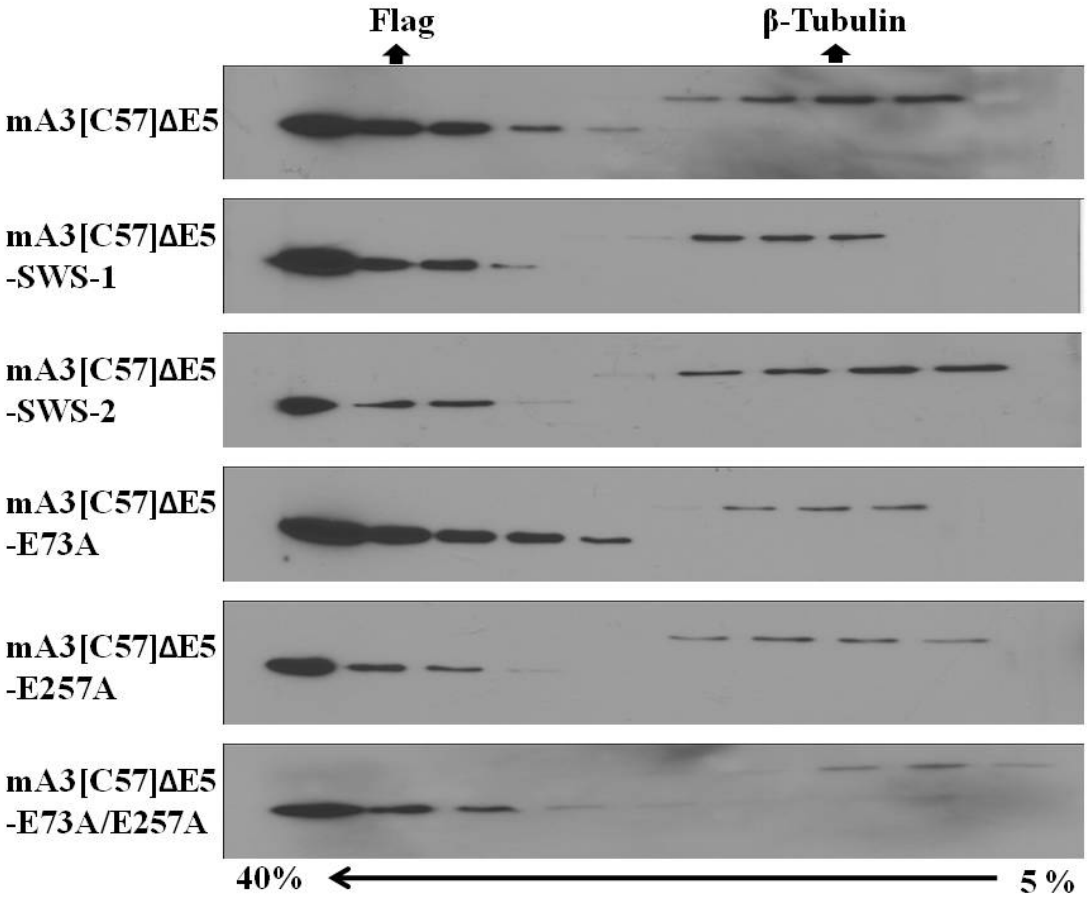

Flag
β-Tubulin
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]
40%
5 %

Figure-34: The molecular mass complex formation of the mA3[C57]ΔE5 protein point mutants *in vitro.* Equal amounts of wild–type mA3[C57]ΔE5 and its point mutant expression vectors (1 μg) were transfected into 293T cells and incubated 48 hours before the velocity sedimentation assay. The mA3[C57]ΔE5 protein point mutants shifted to the bottom of the sucrose gradient in the velocity sedimentation assays showing they are in HMM complexes. Samples were analyzed with anti-Flag western immunoblotting assays in order to determine potential changes in the molecular mass formation. β-Tubulin is a ~55 kDa housekeeping protein, and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.


Flag
β-Tubulin
mA3[C57]ΔE5
mA3[C57]ΔE5
-SWS-1
mA3[C57]ΔE5
-SWS-2
mA3[C57]ΔE5
-E73A
mA3[C57]ΔE5
-E257A
mA3[C57]ΔE5
-E73A/E257A
40%
5 %

5.3 Deamination-dependent and –independent restriction pathways of mA3 and hA3G

My previous infectivity assays revealed that the catalytically inactive mA3 (mA3[C57]ΔE5-E73A/E257A) and hA3G proteins (hA3G-E69A/E259A) still display an antiretroviral activity against the retroviral pathogens during the early stages of infection. In order to elucidate the significance of the deamination-dependent and -independent restriction mechanisms, catalytically inactive mA3[C57]ΔE5-E73A/E257A and hA3G-E69A/E259A proteins were compared with the wild-type proteins. Different experimental strategies have been used to unveil the restrictive capabilities of both the deamination-dependent and –independent restriction pathways of the A3 proteins.

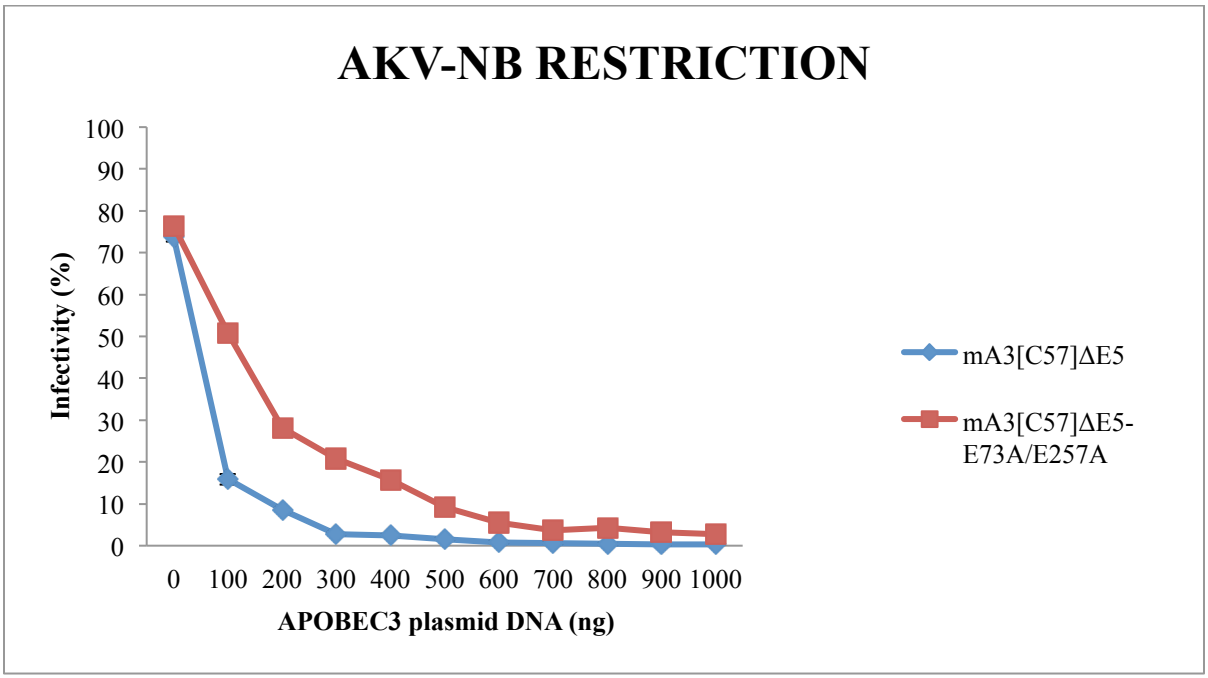
5.3.1 Restriction assays with the enzymatically active and inactive A3 proteins

I initially performed viral restriction assays using the wild-type and catalytically inactive A3 proteins in order to distinguish the difference between the deamination-dependent and –independent restriction pathways of the proteins. Both wild-type and mutated (catalytically inactive) mA3[C57]ΔE5 and hA3G proteins were co-transfected with the AKV-NB virus. The A3 protein concentrations were increased gradually to carefully observe the antiretroviral activity of each mechanism in the restriction assays. The experiments showed that both the hA3G-E69A/E259A and mA3[C57]ΔE5-E73A/E257A catalytically inactive proteins exhibited strong antiretroviral activity against the AKV-NB virus, and this activity was only slightly less potent than their wild-type counterparts, which they exploit both the deamination-dependent and –independent restriction mechanisms (Figure-35A and -35B). In addition, the same experimental approach was also used to target HIVΔvif. This expectedly showed that the catalytically inactive mA3[C57]ΔE5-E73A/E257A and hA3G-E69A/E259A

proteins revealed a weaker antiretroviral activity compared to the AKV-NB virus restriction assay (Figure-36A and -36B). However, there was still a considerable amount of inhibitory effects observed in the experiments with the HIVΔvif virus. I also detected that the deamination-independent restriction mechanism of the catalytically inactive mA3[C57]ΔE5 protein was stronger than the catalytically inactive hA3G protein against the HIVΔvif virus. I have shown that the hA3G protein has the strongest enzymatic activity among the A3 family proteins. Even though the mA3 protein can inhibit the infectivity of HIVΔvif to the same level as the hA3G protein (Figure-36A), the enzymatic activity of the mA3 protein is considerably weaker against the other retroviral pathogens (Table-8). Therefore, I believe it is possible that the mA3 protein may have evolved to acquire a variety of functions to compensate for the weaker enzymatic activity, thereby developing a strong deamination-independent restriction mechanism or other restriction mechanisms that are yet to be determined.

Figure-35: Identification of the deamination-dependent and –independent restriction mechanisms of mA3[C57]ΔE5 and hA3G against AKV-NB by using gradually increasing amounts of A3 in the restriction assays. 1µg viral plasmid DNA was co-transfected with 0 to 1 µg A3 plasmid DNA in the assays. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. A) Comparison of the wild-type (mA3[C57]ΔE5) and the catalytically inactive mA3[C57]ΔE5-E73A/E257A protein in the AKV-NB virus restriction assay. B) Comparison of the wild-type (hA3G) and the catalytically inactive hA3G-E69A/E259A protein in the AKV-NB virus restriction assay.

A)


AKV-NB RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
0 100 200 300 400 500 600 700 800 900 1000
APOBEC3 plasmid DNA (ng)
mA3[C57]ΔE5
mA3[C57]ΔE5-E73A/E257A

B)

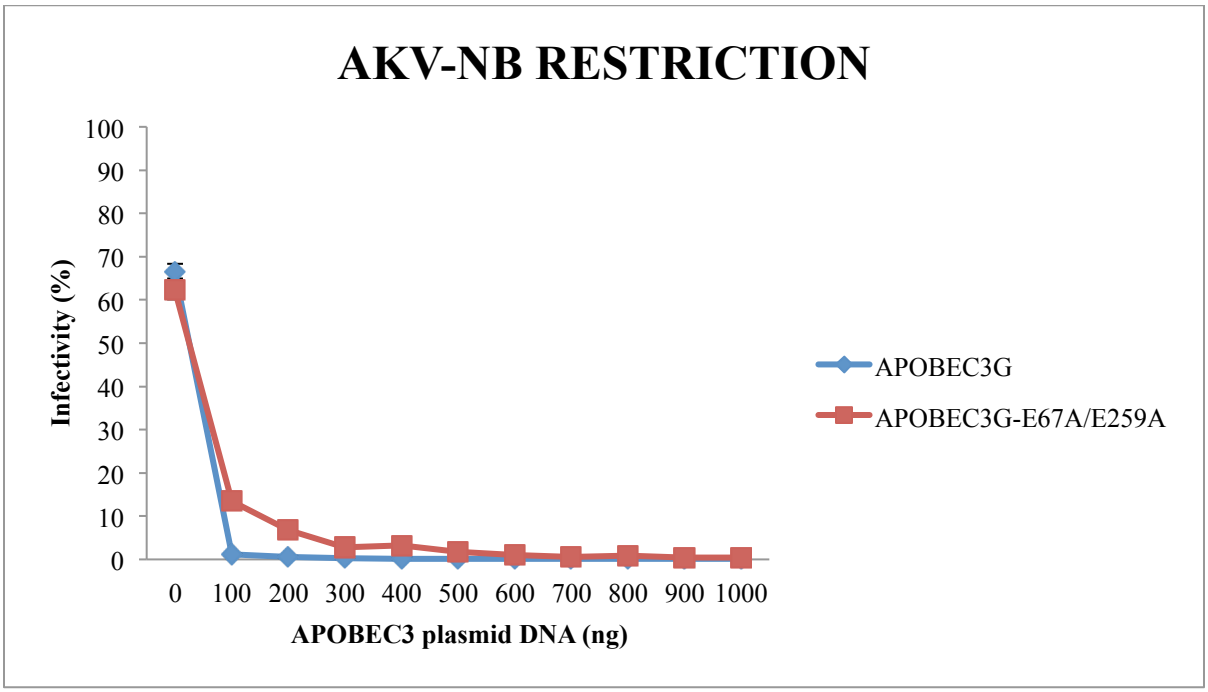
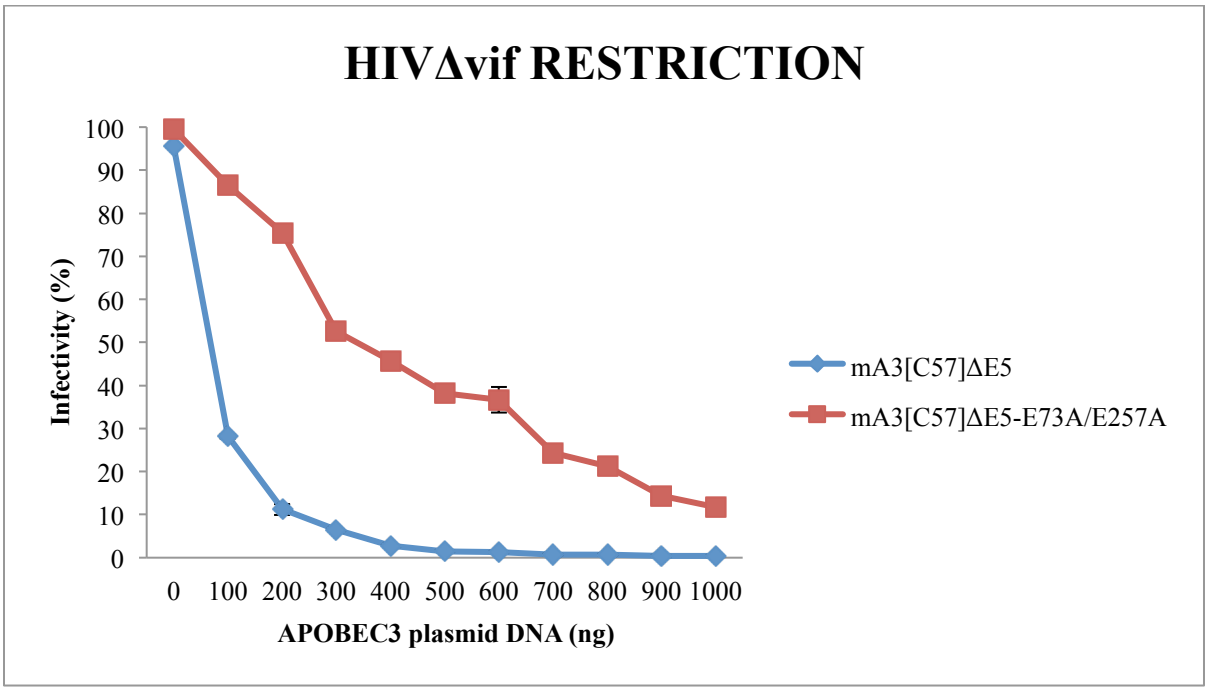
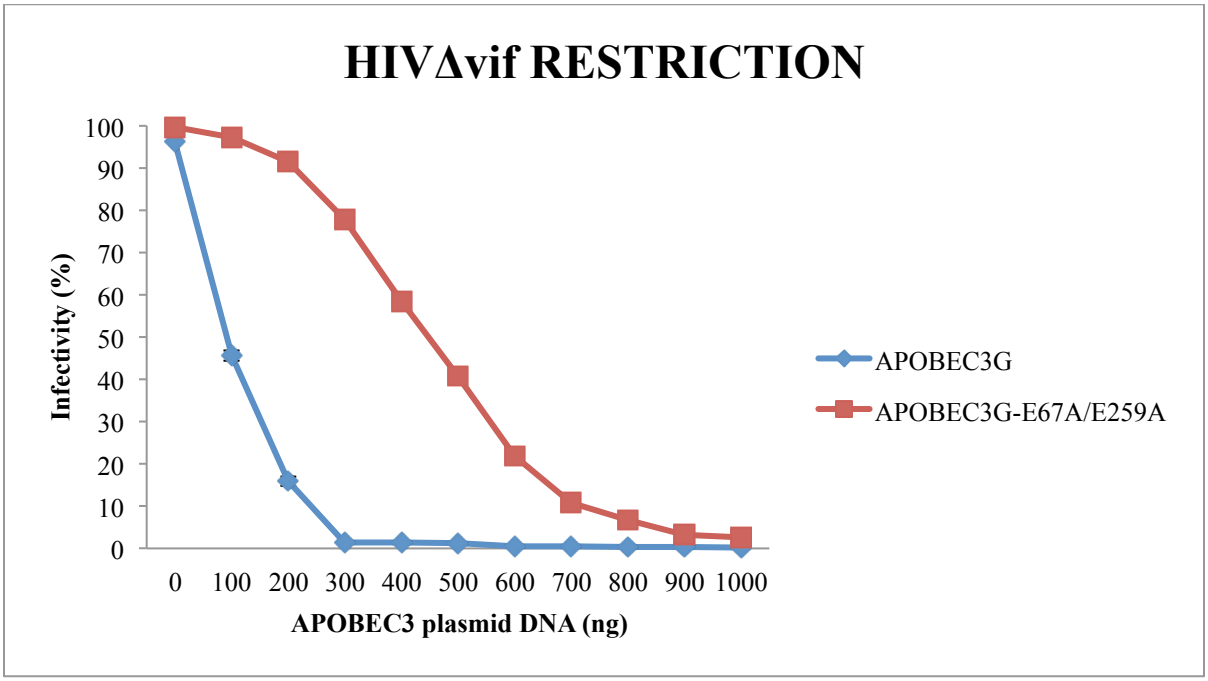

AKV-NB RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
0 100 200 300 400 500 600 700 800 900 1000
APOBEC3 plasmid DNA (ng)
APOBEC3G
APOBEC3G-E67A/E259A

Figure-36: Identification of the deamination-dependent and –independent restriction mechanisms of mA3[C57]ΔE5 and hA3G against HIVΔvif by gradually increasing the concentrations of A3 in the restriction assays. 1μg viral plasmid DNA was co-transfected with 0 to 1 μg A3 plasmid DNA in the assays. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. A) Comparison of the wild-type (mA3[C57]ΔE5) and the catalytically inactive mA3[C57]ΔE5-E73A/E257A protein in the HIVΔvif virus restriction assay. B) Comparison of the wild-type hA3G and the catalytically inactive hA3G-E69A/E259A protein in the HIVΔvif virus restriction assay.

A)


HIVΔvif RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
0 100 200 300 400 500 600 700 800 900 1000
APOBEC3 plasmid DNA (ng)
mA3[C57]ΔE5
mA3[C57]ΔE5-E73A/E257A

B)


HIVΔvif RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
0 100 200 300 400 500 600 700 800 900 1000
APOBEC3 plasmid DNA (ng)
APOBEC3G
APOBEC3G-E67A/E259A

5.3.2 Time course analysis of the deamination-dependent and -independent restriction mechanisms of the mA3 and hA3G protein

In order to investigate the intensity of deamination-independent restriction of mA3[C57]ΔE5 and hA3G proteins, MoMLV, M5P MLV, and AKV-NB viruses were targeted with both the wild-type, hA3G and mA3[C57]ΔE5, and catalytically inactive mA3[C57]ΔE5-E73A/E257A and hA3G-E69A/E259A proteins in a time course experiment. This experiment was designed to measure the potency of the deamination-independent mechanism for a longer period of time than the 24 hours single-cycle infection assays we have been previously conducting. In these experiments, the retrovirus and A3 expression vectors were co-transfected and supernatants of the producer cells were collected to infect target NIH 3T3 cells. Infectivity levels measured by eGFP fluorescence (virus) were analyzed after a 24, 48, and 72 hour incubation time, allowing us to compare restriction between different time points.

As it was previously mentioned, M5P MLV is a single-cycle (self-inactivating) virus, therefore infectivity levels were expected to remain the same in all time points, and the replicative murine gammaretroviruses (AKV-NB and MoMLV) infectivities were expected to increase unless their replication cycle was inhibited by mA3 and hA3G-mediated restriction.

As with my previous results, the time course analysis indicated that 24 hours post-infection, both wild-type mA3[C57]ΔE5 and hA3G proteins inhibited all three viruses to the minimum levels, whereas the catalytically inactive proteins showed weaker but still considerable amount of antiretroviral activity (Figure-37). After 48 hours, the infectivity levels of the non-replicative virus, M5P MLV, remained the same in all samples, and the wild-type A3 protein

restriction also remained at the same level (Figure-38). The infectivity of the replicative viruses, MoMLV and AKV-NB increased over time when produced in the presence of the catalytically inactive A3 proteins (Figure-38). Finally, 72 hours after the infection, the results were astonishingly different than the levels at 24 hours post-infection. The final stage of the time course experiment revealed that the catalytically inactive mutants of mA3[C57]ΔE5 and hA3G were not effective against the replicative viruses as the infectivity levels increased to over 60%, whereas the wild-type proteins were able to keep the infectivity levels of the replicative viruses at the minimum levels throughout the time course analysis (Figure-39).

In conclusion, the results from the three time points indicated that the deamination-independent mechanism has a short-term antiretroviral activity. Without the enzymatic activity of the A3 proteins, viruses tend to escape from the restriction in the target cell. These results clearly show that the replicative viruses can recover from the A3-mediated restriction if both catalytic domains are not active. Therefore, it can be hypothesized that A3 proteins require both deamination-dependent and –independent restriction mechanisms in order to generate complete restriction against the retroviral pathogens. However, the true source of deamination-independent restriction still remains elusive. It should also be mentioned that the main role of the A3 proteins in the innate immune system is to provide an early post-entry block, which these results clearly support. I have shown that the deamination-independent restriction pathway is as important as the deamination-dependent antiviral activity of the A3 proteins during the early stages of infection.

Figure-37: Time course infectivity comparison of the wild-type mA3[C57]ΔE5 and hA3G proteins with the catalytically inactive mutants of these two proteins after 24 hours. The A3 proteins were tested against the AKV-NB, MoMLV and M5P MLV gammaretroviruses and the infectivity levels were measured by flow cytometry analysis. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population.

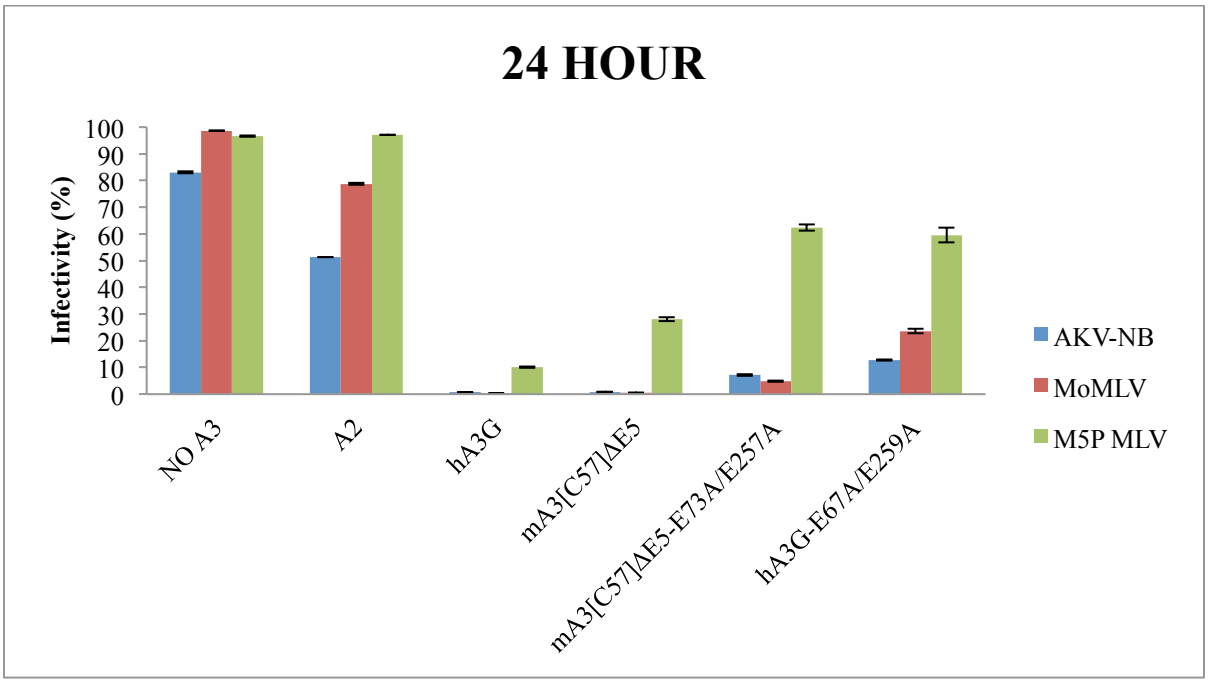

24 HOUR
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-E73A/E257A
hA3G-E67A/E259A
AKV-NB
MoMLV
M5P MLV

Figure-38: Time course infectivity comparison of the wild-type mA3[C57]ΔE5 and hA3G proteins with the catalytically inactive mutants of these two proteins after 48 hours. The A3 proteins were tested against the AKV-NB, MoMLV and M5P MLV gammaretroviruses and the infectivity levels were measured by flow cytometry analysis. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population.

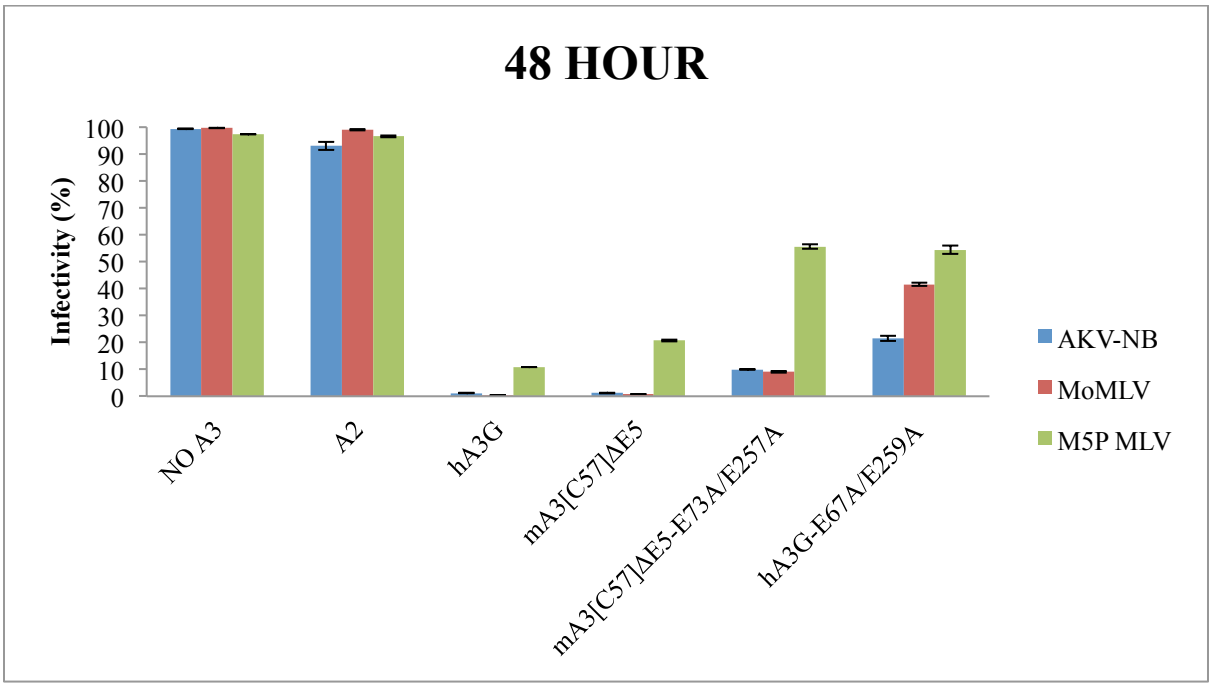
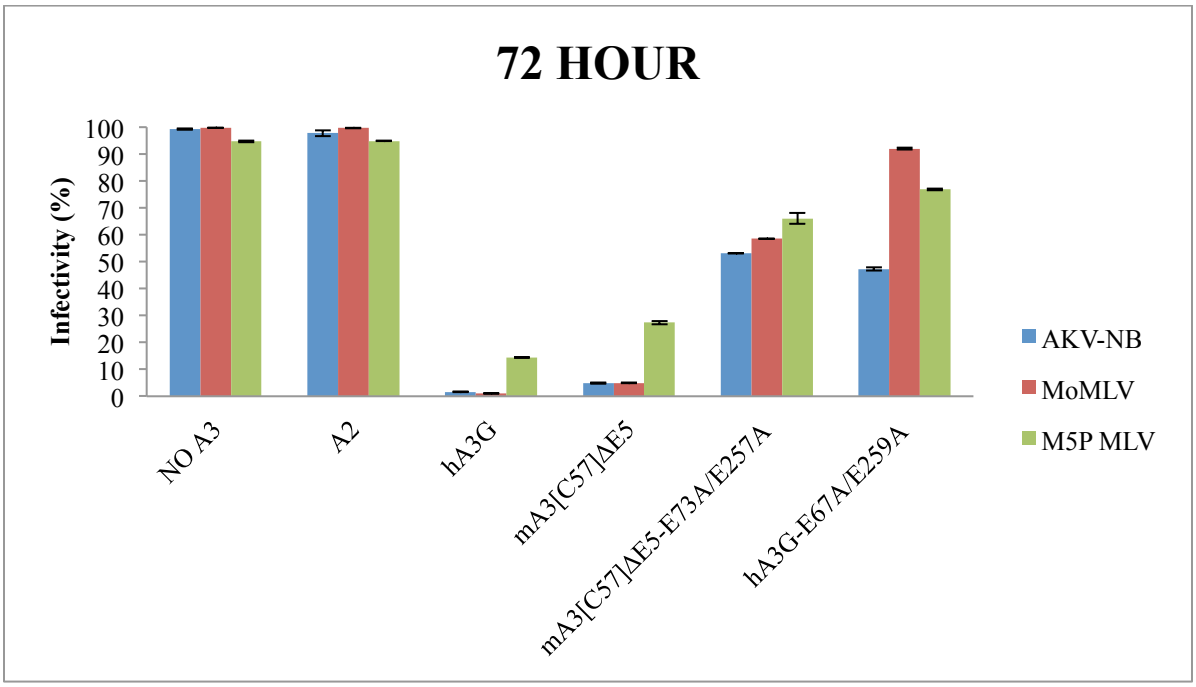

48 HOUR
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-E73A/E257A
hA3G-E67A/E259A
AKV-NB
MoMLV
M5P MLV

Figure-39: Time course infectivity comparison of the wild-type mA3[C57]ΔE5 and hA3G proteins with the catalytically inactive mutants of these two proteins after 72 hours. The A3 proteins were tested against the AKV-NB, MoMLV and M5P MLV gammaretroviruses and the infectivity levels were measured by flow cytometry analysis. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population.

.


72 HOUR
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-E73A/E257A
hA3G-E67A/E259A
AKV-NB
MoMLV
M5P MLV

5.3.3 Mutation analysis for the direct detection of mA3 and hA3G-mediated single-stranded DNA editing

DNA editing is one of the main retroviral restriction mechanisms of mA3 and the detection of G-to-A transition mutations in the integrated retroviral genome is a direct approach to determine the enzymatic activity of each catalytic domain of the mA3 protein. As previously discussed, the mA3 protein can mutate HIV-1 single-stranded DNA replication intermediates as shown by the detection of high numbers of mutations in HIV-1 integrated sequences.

To further investigate the enzymatic activity of mA3, infectivity assays were performed by producing HIVΔvif in the presence of the four mA3 splicing isoforms. Twenty four hours post-infection, genomic 293T cell DNA was extracted. PCR analysis was then performed to amplify the target eGFP region in the viral genome. Amplification of the correct region was determined by gel electrophoresis, followed by cloning and sequence analysis for G to A mutations.

5.3.3.1 Mutation frequency of the mA3 splicing isoforms on HIVΔvif

My analysis indicates that all the mA3 splicing isoforms, except the low expresser mA3[Balb], caused approximately 15 mutations per kb, and this ratio was similar to the mutation rate of the hA3G protein against HIVΔvif (Table-3). These results confirm that three out of four mA3 splicing isoforms elicit a high mutation rate against the HIVΔvif virus. Previous infectivity assays revealed that mA3[C57] does not fully restricts the HIVΔvif lentivirus but decreases the infectivity levels partially (Figure-25). However, mutation analysis indicates that this protein has high enzymatic activity and generates approximately 15 mutations per kb on viral DNA sequences. Therefore, it was hypothesized that the

deamination-independent pathway of this protein may not be effective against HIVΔvif lentivirus. In addition, infectivity assays with mA3[Balb] and HIVΔvif also resulted in the partial restriction of the virus but subsequent sequencing analysis detected 5.8 mutation per kb on viral sequences. As a result, it was concluded that mA3 protein with exon 5 may have reduced deamination-independent activity against HIVΔvif but the lower restriction with mA3[Balb] may also be linked to low expression levels of this protein.

During the mutation analysis with HIVΔvif and other retroviruses, sequences containing G to A mutations were scrutinized in order to select the mutations occurring in an A3 context. It is known that reverse transcriptase of these viruses does not have any proof-reading activity and sequence errors may also originate from PCR assays or sequencing analysis. Therefore, in order to exclude these issues, mutation specificity of each sequence was analyzed very carefully and only mutations with A3 context (e.g.: cC, tC) were added to the data analysis.

Table-3: Determination of the enzymatic activity of the mA3 splicing isoforms and the hA3G protein against HIVΔvif. Target cells were infected and later harvested. DNA was isolated and sequenced for G-to-A transition mutations in the integrated retroviral genome (eGFP region). The number of bp sequenced, percentage of mutated sequences, number of G-to-A mutations and the mutation rate per kb were calculated for each mA3 protein splicing isoform and for the hA3G protein. Spurious sequences were expelled from the dataset prior to alignment and mutation analysis.

HIV-1 Mutation Analysis	Number of bp sequenced	Percentage of sequences mutated (%)	Number of G to A mutations	Mutation rate (mutation/kb)
hA3G	3360	100	51	15.2
mA3[C57]	6240	92	88	14.1
mA3[C57]ΔE5	11510	85	179	15.5
mA3[Balb]ΔE5	11440	76	189	16.5
mA3[Balb]	6560	50	38	5.8

5.3.3.2 Enzymatic activity of mA3[C57]ΔE5 and point mutants of the protein on HIVΔvif

Examination of the enzymatic activity of each of the mA3 catalytic domains was also performed on the HIVΔvif virus. My results show that only one of the catalytic domains of the mA3 protein is enzymatically active, as well, the domain organization of the mA3 protein is reversed compared to the hA3G protein. In our experiments, the point mutation in the first catalytic domain (mA3[C57]ΔE5-E73A) abolished the deaminase activity of the protein entirely. In contrast, the mA3 protein with a point mutation in the second domain (mA3[C57]ΔE5-E257A) was still enzymatically active, even though this activity was almost 50% less than the wild-type protein. Considering the second domain has no detectable enzymatic activity, perhaps the second domain facilitates the enzymatic activity of the first domain. It may also be involved in other mechanisms related with the deamination-independent restriction pathway of the protein that has yet to be determined (Table-4). Unfortunately, the reason behind the reduced enzymatic activity of the second domain mutant is still unknown, but it is an important issue that needs to be solved in order to reveal the restriction mechanisms of the protein. The catalytically inactive mA3 protein with both of these point mutations (mA3[C57]ΔE5-E73A/E257A) was also enzymatically inactive.

Table-4: Infectivity assays to determine the enzymatic activity of the catalytic domains of mA3 by producing HIVΔvif with mA3[C57]ΔE5 point mutants. Number of bp sequenced, percentage of mutated sequences, number of G-to-A mutations and mutation rate per kb were calculated for each point mutant and the wild-type mA3[C57]ΔE5 protein. Spurious sequences were expelled from the dataset prior to alignment and mutation analysis.

HIV-1 Mutation Analysis	Number of bp sequenced	Percentage of sequences mutated (%)	Number of G to A mutations	Mutation rate (mutation/kb)
mA3[C57]ΔE5	11510	85	179	15.5
mA3[C57]ΔE5-E73A	4010	-	-	-
mA3[C57]ΔE5-E257A	7810	92	65	8.3
mA3[C57]ΔE5-E73A/E257A	2660	-	-	-

5.3.3.3 Deamination DNA context specificity of the mA3 splicing isoforms and the mA3[C57]ΔE5 point mutants

In order to validate that the enzymatic activity is a result of mA3 and hA3G deamination, I investigated the sequence context in which deamination is expected to occur. It was known from previous reports that the mA3 protein targets the third cytidine in a 3' TT**C** 5' or 3' TC**C** 5' context on the minus strand DNA intermediates and converts them to uridine. The hA3G protein also targets the third cytidine in a 3' TC**C** 5' or 3' CC**C** 5' context. The mutation specificity of these proteins on the plus strand DNA was analyzed for each mutation we discovered within the eGFP gene. In conclusion, I identified that the mutations in the HIVΔvif retroviral genome were indeed in the sequence contexts mentioned above, and the mutation specificity for each mA3 splicing isoform was described (Table-5, -6, -7). Regarding their sequence specificity, there were no major differences among the mA3 splicing isoforms but the mA3 mutations were clearly distinguished from hA3G mutations during the sequence analysis.

Table-5: Mutation specificity of the hA3G protein and the mA3[C57] splicing isoforms on HIVΔvif integrated genomic DNA. Mutations were detected in the eGFP gene of HIVΔvif when the pseudoviruses were produced in presence of hA3G (A), mA3[C57]ΔE5 (B), and mA3[C57] (C). As previously described, A3 proteins convert cytidines into uridines on the minus stand DNA, but mutations analysis in my experiments were performed on the plus strand DNA in order to observe the G-to-A transitions on the integrated proviral genomes. The G-to-A mutations detected on the plus DNA strand were considered point zero and mutation specificity was examined by screening towards 3' direction because A3 proteins deaminate from 3' to 5' direction. The base on the right side of the mutation was considered as -1 position, and the one next to it was identified as -2 position. These three bases were then plotted on the table in order to identify the mutation specificity of the protein. Tables were prepared regarding the plus strand observations and the minus strand conversion of mutation specificity was not indicated.

A)

HIV + hA3G	0	-1	-2
A	-	14	25
C	-	-	12
T	-	4	14
G	100	82	49

B)

HIV + mA3[C57]ΔE5	0	-1	-2
A	-	38	62
C	-	22	6
T	-	3	17
G	100	37	15

C)

HIV + mA3[C57]	0	-1	-2
A	-	59	76
C	-	3	9
T	-	1	5
G	100	37	10

Table-6: Mutation specificity of the mA3[Balb] splicing isoforms on HIVΔvif . Mutations were detected in the eGFP gene of integrated HIVΔvif genomic DNA when pseudoviruses were produced in presence of mA3[Balb]ΔE5 (A), and mA3[Balb] (B). As previously described, A3 proteins convert cytidines into uridines on the minus stand DNA, but mutations analysis in my experiments were performed on the plus strand DNA in order to observe the G-to-A transitions on the integrated proviral genomes. The G-to-A mutations detected on the plus DNA strand were considered point zero and mutation specificity was examined by screening towards 3' direction because A3 proteins deaminate from 3' to 5' direction. The base on the right side of the mutation was considered as -1 position, and the one next to it was identified as -2 position. These three bases were then plotted on the table in order to identify the mutation specificity of the protein. Tables were prepared regarding the plus strand observations and the minus strand conversion of mutation specificity was not indicated.

A)

HIV + mA3[Balb]ΔE5	0	-1	-2
A	-	61	74.
C	-	10	8
T	-	2	4
G	100	27	14

B)

HIV + mA3[Balb]	0	-1	-2
A	-	63	76
C	-	8	6
T	-	-	2
G	100	29	16

Table-7: Mutation specificity of the mA3[C57]ΔE5 point mutants were investigated on HIVΔvif . Mutations were detected for mA3[C57]ΔE5-E257A (B) but not with other mutants (A and C). Mutations were detected in the eGFP gene of HIVΔvif when the pseudovirus was produced in presence of mA3[C57]ΔE5. As previously described, A3 proteins convert cytidines into uridines on the minus stand DNA, but mutations analysis in my experiments were performed on the plus strand DNA in order to observe the G-to-A transitions on the integrated proviral genomes. The G-to-A mutations detected on the plus DNA strand were considered point zero and mutation specificity was examined by screening towards 3' direction because A3 proteins deaminate from 3' to 5' direction. The base on the right side of the mutation was considered as -1 position, and the one next to it was identified as -2 position. These three bases were then plotted on the table in order to identify the mutation specificity of the protein. Tables were prepared regarding the plus strand observations and the minus strand conversion of mutation specificity was not indicated.

A)

HIV + mA3[C57]ΔE5-E73A	0	-1	-2
A	-	-	-
C	-	-	-
T	-	-	-
G	-	-	-

B)

HIV + mA3[C57]ΔE5-E257A	0	-1	-2
A	-	38	72
C	-	19	9
T	-	1	11
G	100	42	8

C)

HIV + mA3[C57]ΔE5-E73A/E257A	0	-1	-2
A	-	-	-
C	-	-	-
T	-	-	-
G	-	-	-

5.4 Analysis on the potential role of glycosylation of the glycogag protein on gammaretrovirus restriction by mA3:

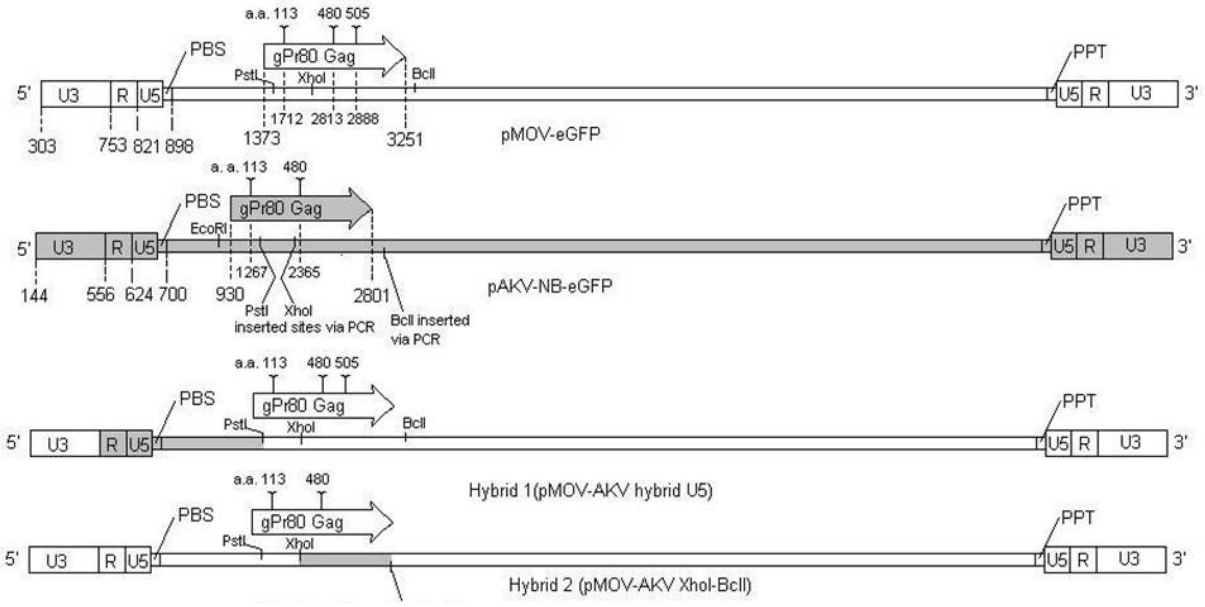
Glycogag is a glycosylated form of the gag polyprotein that initiates at an alternative start codon located upstream of the AUG start codon of the gag polyprotein. Glycogag contains an additional 88 amino acids, which includes an ER localization signal peptide (41, 46, 55). In addition, gammaretroviruses also contain three glycosylation sites in the gag polyprotein. The glycosylation sites are within the gag polyprotein; however the function of these sites is largely unknown. There is some evidence that the glycogag protein may be important for productive infection *in vivo* (55). Because murine gammaretroviruses do not encode an accessory gene such as *Vif,* which can inhibit the antiretroviral activity of the mA3 protein, it was hypothesized that the glycogag protein of these viruses could similarly affect the infectivity of the gammaretroviruses (32).

Here, hybrid gammaretroviruses were generated to investigate whether these regions interact with mA3 during the retroviral infectious cycle and potentially interfere with the enzymatic activity of the protein. Both Hybrid-1 MLV and Hybrid-2 MLV contain MoMLV retroviral sequences, but the regions encompassing the leader sequence of the glycogag protein (Hybrid-1 MLV) and the internal glycosylation sites (Hybrid-2 MLV) of these gammaretroviruses are that of AKV-NB which is sensitive to restriction and deamination by mA3. We chose AKV-NB virus to generate these hybrids because studies indicated that there are detectable mutations in the AKV-NB genome but not in the MoMLV sequences. It was therefore believed that glycosylation might be a protective mechanism against the enzymatic activity of the mA3 protein

Figure-40: Schematic representation of the glycogag hybrid gammaretroviruses. Hybrid-1 MLV was generated to assess the importance of the glycogag segment. The region from the end of 5' U3 segment to the gPr65 gag AUG start codon was swapped with the same region in the AKV-NB virus. Also, MoMLV virus contains three glycosylation sites at amino acid (aa positions) 113, 480, and 505, respectively. The region containing all three glycosylation sites was swapped with AKV-NB virus to generate the Hybrid-2 MLV gammaretrovirus.

(MoMLV)

(AKV-NB)


a.a. 113
480 505
PBS
gPr80 Gag
PstI
XhoI
BclI
PPT
5'
U3
R
U5
U5
R
U3
3'
303
753
821
898
1373
1712
2813
2888
3251
pMOV-eGFP
a. a. 113
480
EcoRI
144
556
624
700
930
1267
2365
2801
PstI
XhoI
inserted sites via PCR
BclI inserted
via PCR
pAKV-NB-eGFP
Hybrid 1(pMOV-AKV hybrid U5)
Hybrid 2 (pMOV-AKV XhoI-BclI)

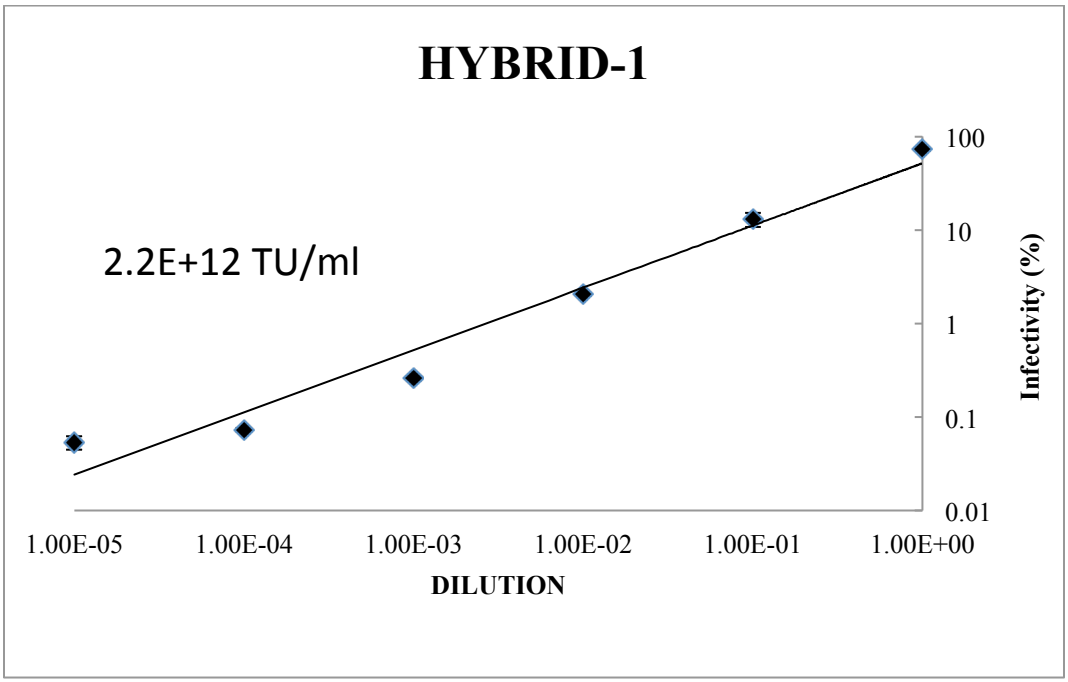
5.4.1 Titration analysis of the replicative Hybrid MLV gammaretroviruses:

Titration analysis for Hybrid-1 MLV and Hybrid-2 MLV were performed to investigate the *in vitro* infectivity levels of these viruses, and also to compare the activity of these genetically engineered viruses to the original gammaretroviruses, AKV-NB and MoMLV. The infectivity levels of the hybrid viruses were measured in a serial tenfold dilution system and transducing units (TU/ml) were calculated (Figure-41). Titration analysis indicated that the two hybrid gammaretroviruses have different infectivity levels. Interestingly, Hybrid-1 MLV had similar infectivity levels with AKV-NB virus and Hybrid-2 revealed similar results to MoMLV. This suggests that it was not the glycosylation directly, but the leader sequence and the glycogag protein of the gammaretroviruses that could play a role in the viral kinetics of infectivity (Figure-42, -43).

Furthermore, the interactions between the hybrid gammaretroviruses and the mA3 splicing isoforms were also examined to understand whether the glycosylation of the gag protein had an effect on mA3-mediated restriction. The infectivity assays show that the mA3 splicing isoforms are strongly active against both hybrid gammaretroviruses during the early stages of infection (24 hours) and the restriction levels were similar to the original AKV-NB and MoMLV gammaretroviruses. In conclusion, infectivity assays with Hybrid MLVs also showed that the strong deamination-independent restriction mechanism may heavily contribute to the antiretroviral activity of the mA3 splicing isoforms during the early stages of infection. Further analysis is necessary to determine if there is a link between the enzymatic activity of mA3 and the glycosylation of the gag protein of the hybrid gammaretroviruse

Figure-41: Titration Analysis of the replicative Hybrid-1 MLV and Hybrid-2 MLV mouse gammaretroviruses. Serial tenfold dilutions of the hybrid gammaretroviruses were used to infect target NIH 3T3 cells. The infectivity levels were measured 24 hours after the spin-infection. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. A) The infectivity levels of the Hybrid-1 MLV gammaretrovirus (TU/ml) and the titration analysis for each serial dilution. B) The infectivity levels of the Hybrid-2 MLV gammaretrovirus (TU/ml) and the titration analysis for each serial dilution.

A)


HYBRID-1
2.2E+12 TU/ml
100
10
1
0.1
0.01
Infectivity (%)
1.00E-05
1.00E-04
1.00E-03
1.00E-02
1.00E-01
1.00E+00
DILUTION

B)

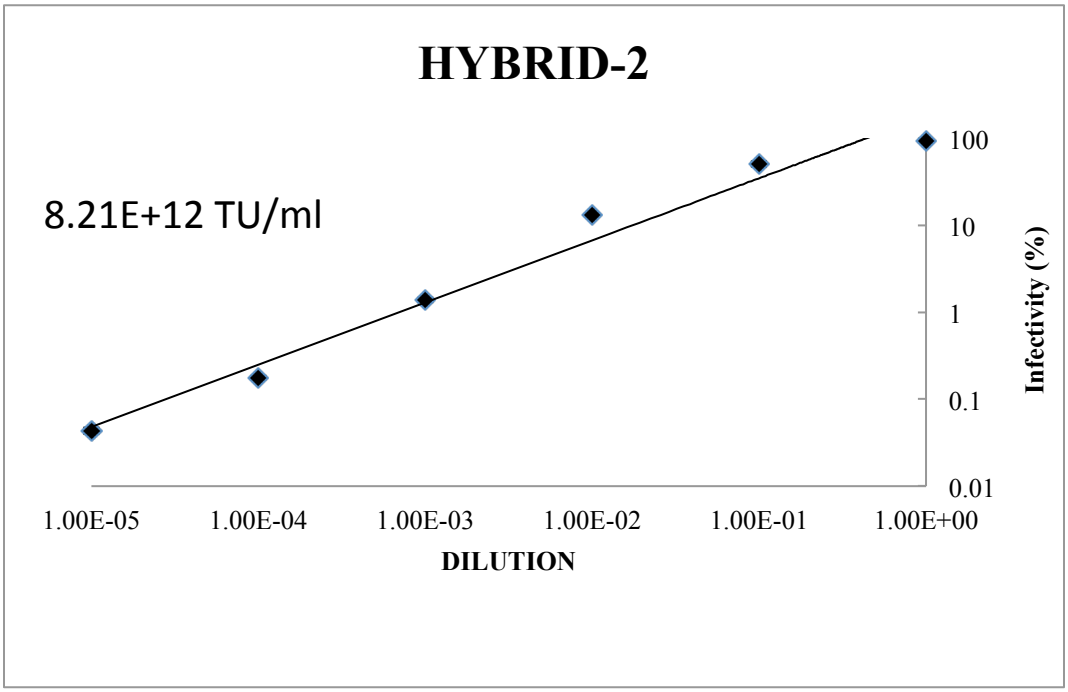
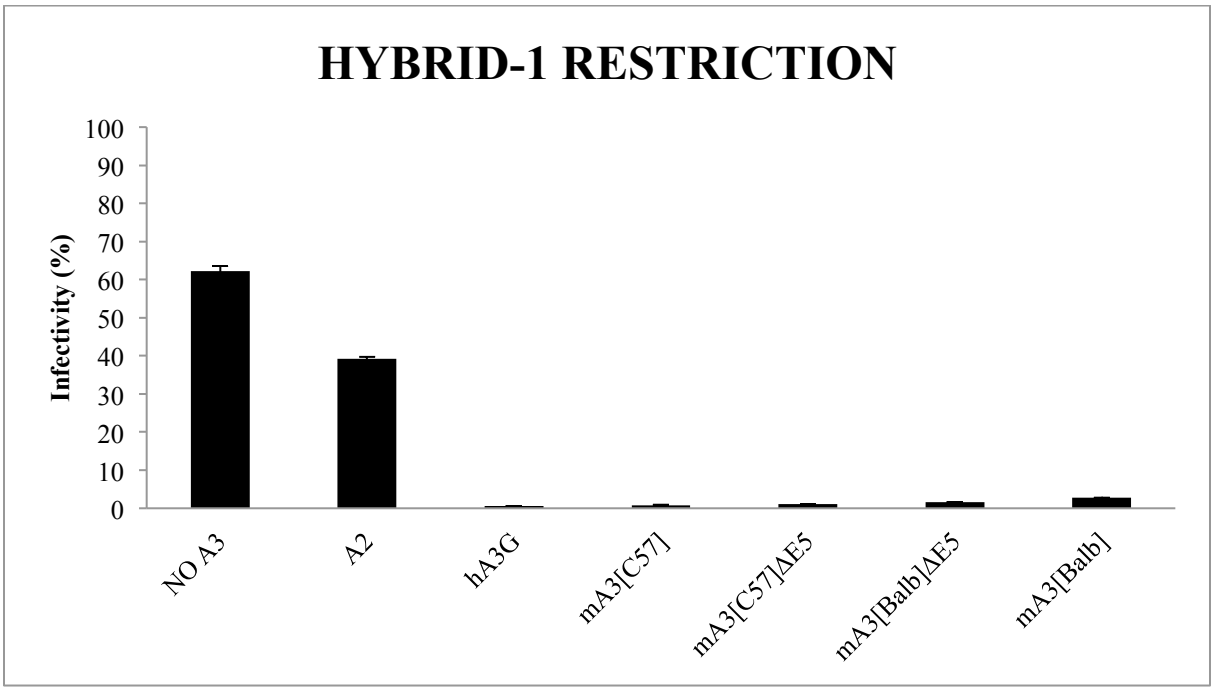

HYBRID-2
8.21E+12 TU/ml
100
10
1
0.1
0.01
Infectivity (%)
1.00E-05
1.00E-04
1.00E-03
1.00E-02
1.00E-01
1.00E+00
DILUTION

Figure-42: Antiretroviral activity of the mA3 splicing isoforms against Hybrid-1 MLV *in vitro*. In order to assess the restriction at the early stages of infection, 25 ng of A2, 100 ng of hA3G and 300 ng of each mA3 splicing isoform were co-transfected with 1µg viral plasmid DNA in 293T cells. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population.A) mA3-mediated restriction of the Hybrid-1 MLV gammaretrovirus was observed after a single-cycle infection. B) Protein expression levels of the control samples and the mA3 splicing isoforms were measured in 293T producer cells with an anti-Flag western immunoblotting assay. β-Tubulin is a ~55 kDa housekeeping protein, and full-length Flag-mA3 proteins are ~51 kDa and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


HYBRID-1 RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]

B)

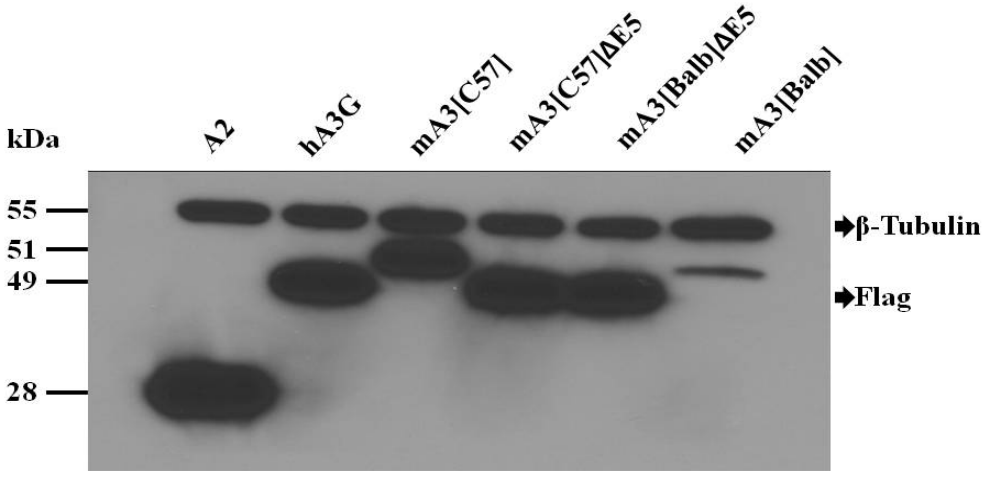
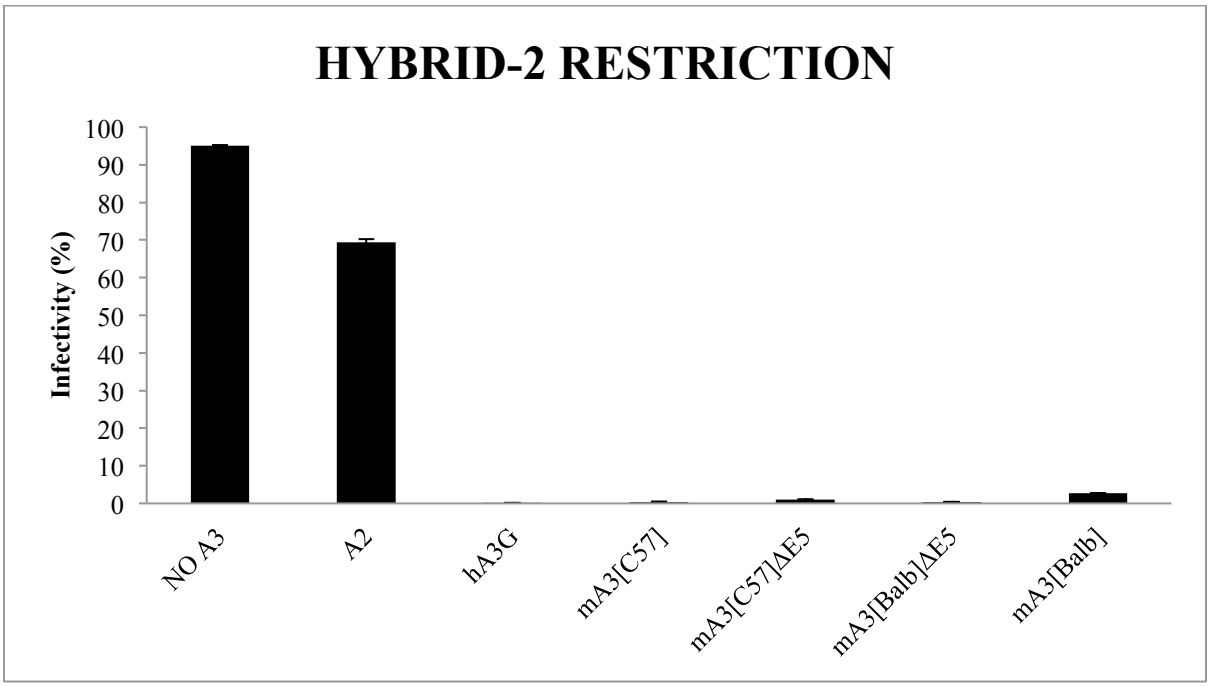
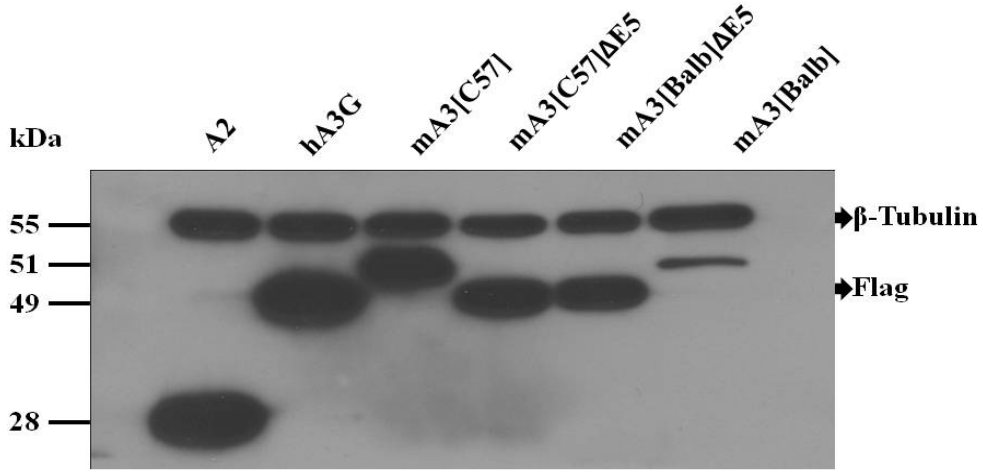

kDa
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]
55
51
49
28
➡β-Tubulin
➡Flag

Figure-43: Antiretroviral activity of mA3 splicing isoforms against Hybrid-2 MLV *in vitro*. In order to assess the restriction at the early stages of infection, 25 ng of A2, 100 ng of hA3G and 300 ng of each mA3 splicing isoform were co-transfected with 1µg viral plasmid DNA in 293T cells. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population. A) mA3-mediated restriction of the Hybrid-2 MLV gammaretrovirus was observed after a single-cycle infection. B) Protein expression levels of the control samples and mA3 splicing isoforms were measured in 293T producer cells with anti-Flag western immunoblotting assay. β-Tubulin is a ~55 kDa housekeeping protein, and full-length Flag-mA3 proteins are ~51 kDa and Flag-mA3 proteins lacking exon 5 are ~49 kDa size.

A)


HYBRID-2 RESTRICTION
Infectivity (%)
100
90
80
70
60
50
40
30
20
10
0
NO A3
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]

B)


kDa
A2
hA3G
mA3[C57]
mA3[C57]ΔE5
mA3[Balb]ΔE5
mA3[Balb]
55
51
49
28
β-Tubulin
Flag

5.4.2 Enzymatic activity of mA3[C57]ΔE5 against murine gammaretroviruses:

The enzymatic activity of mA3[C57]ΔE5 was also tested against the murine gammaretroviruses and this activity was compared with the hA3G protein in the experiments. In addition, it was also critical to investigate whether glycosylation of gag hindered the deamination-dependent mechanisms of mA3. MoMLV, Hybrid-1 MLV, Hybrid-2 MLV and AKV-NB viral sequences were analyzed for their susceptibility to the enzymatic activity of mA3[C57]ΔE5. My sequence analysis indicate that mA3[C57]ΔE5 was only enzymatically active against AKV-NB and there were no detectable mutations in the integrated genomes of MoMLV and the hybrid gammaretroviruses. In contrast, hA3G protein had strong deaminase activity against MoMLV, and mutation rate was almost similar to the mutation rate on HIVΔvif (Table-8).

In summary, the enzymatic activity of the mA3[C57]ΔE5 protein was not detectable against MoMLV and hybrids, and only the AKV-NB genome had detectable mutations. The MoMLV and hybrid viruses protected themselves from the enzymatic activity of the mA3 protein by an unidentified mechanism. However, mA3 splicing isoforms were able to inhibit the infectivity of gammaretroviruses after a single cycle of infection (24 hours), and this activity was attributed to the deamination-independent restriction mechanisms of the protein. Moreover, mutation specificity was also analyzed from mA3[C57]ΔE5 and the hA3G proteins (Table-9, -10). Similarly, analysis indicated that the G-to-A transitions detected on the plus strand MoMLV sequences were generated by hA3G protein. However, after investigating more than 100 different clones in the assays, no detectable mA3[C57]ΔE5 mutations have been found on MoMLV, Hybrid-1 MLV, and Hybrid-2 MLV sequences.

Table-8: Determination of the enzymatic activity of mA3[C57]ΔE5 and hA3G on mouse gammaretroviruses. Genomic DNA from the NIH 3T3 target cells in my infectivity assays was extracted and analyzed for G-to-A transition mutations in the integrated retroviral DNA. MoMLV, Hybrid-1 MLV, Hybrid-2 MLV, and AKV-NB viruses were used in the experiments to examine the relationship between glycosylation and mA3 enzymatic activity. The number of bp sequenced, the percentage of mutated sequences, the number of G-to-A transition mutations and the mutation rate per kb were calculated for mA3[C57]ΔE5 and hA3G. Spurious sequences were expelled from the dataset prior to alignment and mutation analysis.

MLV Mutation Analysis	Number of bp sequenced	Percentage of sequences mutated (%)	Number of G to A mutations	Mutation rate (mutation/kb)
MoMLV + hA3G	6390	77.8	77	12.05
MoMLV + mA3[C57]ΔE5	13440	0	0	0
HYBRID-1 + mA3[C57]ΔE5	9700	0	0	0
HYBRID-2 + mA3[C57]ΔE5	12950	0	0	0
AKV-NB + mA3[C57]ΔE5	7150	58.4	12	1.8

Table-9: Mutation specificity of hA3G and mA3[C57]ΔE5 against murine gammaretrovirus DNA. Mutations were investigated in the eGFP gene of MoMLV and AKV-NB integrated DNA sequences. A) Mutation specificity of hA3G against MoMLV DNA. B) No detectable mutations were induced by mA3[C57]ΔE5 on MoMLV DNA. C) Mutation specificity of mA3[C57]ΔE5 against the DNA of the AKV-NB virus. As previously described, A3 proteins convert cytidines into uridines on the minus stand DNA, but mutations analysis in our experiments were performed on the plus strand DNA in order to observe the G-to-A transitions on the integrated proviral genomes. The G-to-A mutations detected on the plus DNA strand were considered point zero and mutation specificity was examined by screening towards 3' direction because A3 proteins deaminate from 3' to 5' direction. The base on the right side of the mutation was considered as -1 position, and the one next to it was identified as -2 position. These three bases were then plotted on the table in order to identify the mutation specificity of the protein. Tables were prepared regarding the plus strand observations and the minus strand conversion of mutation specificity was not indicated.

A)

MoMLV + hA3G	0	-1	-2
A	-	14	28
C	-	3	7
T	-	-	15
G	100	83	50

B)

MoMLV + mA3[C57]ΔE5	0	-1	-2
A	-	-	-
C	-	-	-
T	-	-	-
G	-	-	-

C)

AKV-NB + mA3[C57]ΔE5	0	-1	-2
A	-	25	50
C	-	17	8
T	-	33	17
G	100	25	25

Table-10: Mutation analysis of mA3[C57]ΔE5 against hybrid MoMLV gammaretroviruses. Mutations were investigated in the eGFP gene of the hybrid MLV integrated retroviral DNA. No detectable deaminase activity by the mA3[C57]ΔE5 protein was seen against Hybrid-1 MLV (A) and Hybrid-2 MLV gammaretroviruses (B). As previously described, A3 proteins convert cytidines into uridines on the minus stand DNA, but mutations analysis in our experiments were performed on the plus strand DNA in order to observe the G-to-A transitions on the integrated proviral genomes. The G-to-A mutations detected on the plus DNA strand were considered point zero and mutation specificity was examined by screening towards 3' direction because A3 proteins deaminate from 3' to 5' direction. The base on the right side of the mutation was considered as -1 position, and the one next to it was identified as -2 position. These three bases were then plotted on the table in order to identify the mutation specificity of the protein. Tables were prepared regarding the plus strand observations and the minus strand conversion of mutation specificity was not indicated.

A)

HYBRID-1 + mA3[C57]ΔE5	0	-1	-2
A	-	-	-
C	-	-	-
T	-	-	-
G	-	-	-

B)

HYBRID-2 + mA3[C57]ΔE5	0	-1	-2
A	-	-	-
C	-	-	-
T	-	-	-
G	-	-	-

6.0 DISCUSSION

Using mA3 splicing isoforms, mA3[C57]ΔE5 point mutants, and HIVΔvif and murine gammaretroviruses, the structural organization of the mA3 protein and the antiretroviral function of the each catalytic domain of the mA3 intrinsic antiretroviral restriction factor was investigated in this study.

Firstly, it is important to note that the assays in the study were carefully designed to analyze the restriction mechanisms of the mA3 protein at the early stages of infection. The aim of the study was to understand how the mA3 protein works *in vitro* to restrict retroviruses. In order to use the knowledge gained from this work for subsequent *in vivo* studies on retroviral pathogens, efforts were made to conduct our experiments in such ways as to reduce artefacts generated from conducting our experiments in non-physiological conditions. For instance, I optimized my infection systems to use the minimal amounts of APOBEC3 and virus that would allow for reliable characterization, while still providing consistent results. As such, the experimental strategy was designed to obtain the most accurate answers possible to understand the endogenous function of the mA3 protein.

Specifically, the first objective of the project was to analyze the antiretroviral activity of the mA3 protein splicing isoforms and to identify the functional role of each of the catalytic domains against HIVΔvif and murine gammaretroviruses *in vitro*. Initially, the expression levels of the mA3 splicing isoforms were investigated in a time course experiment where it was determined that protein expression levels were at their peak levels after a 48 hour incubation period (Figure-10A, -10B). Also, the results revealed that the mA3 splicing isoforms had similar transfection efficiencies compared with hA3G (Figure-10A). However,

the mA3[Balb] isoform revealed low expression levels by FACS analysis and by WB (Figure-10B). In my study, mA3[Balb] consistently showed the lowest expression levels and normalization among the samples was required to allow for similar expression levels of the full length mA3[Balb] allele compared to the other mA3 splicing isoforms (Figure-11). The mA3[Balb]ΔE5 isoform also had similar transfection and expression levels compared to the other mA3[C57] splicing isoforms at 24 hours but further time point analysis (48, 72, 96 hours) indicated that this isoform tends to deplete faster from our extracts than its human and C57BL/6 mouse counterparts *in vitro* (Figure-10B). Moreover, point mutations in mA3[C57]ΔE5 did not affect protein expression levels of these mutants when measured at 48 hours (Figure-12).

In order to design a quantifiable *in vitro* experimental model to investigate the restriction of retroviral pathogens, the infectivity levels of the laboratory viruses were also standardized. I carefully calculated the titers of each virus produced by standard transfection conditions. (Figure-13, -14, -15, -41).

After determining the expression levels of the mA3 splicing isoforms in my experiments, restriction assays were subsequently performed to investigate the antiretroviral activities of these proteins against the various retroviral pathogens. First, increasing amounts of mA3 were used to identify the optimal restriction point for future infectivity assays (Figure-17). The antiretroviral activity of the mA3[C57]ΔE5 protein was tested on various retroviral pathogens and it was concluded that 300 ng of mA3 plasmid DNA is enough to impair the infectious cycle of all retroviruses tested, except M5P MLV (Figure-16). Also, the same experimental approach was used to analyze how my negative control A2 behaved against the AKV-NB virus (Figure-18) and how effective my positive control, hA3G, could restrict the

AKV-NB, M5P MLV and HIVΔvif viruses (Figure-17, -19). In conclusion, the first part of my work served to establish the optimal conditions to carry out the experiments that will help dissect what structural components of the protein are involved in the retroviral restriction.

Retroviral restriction studies revealed that all mA3 splicing isoforms have a strong inhibitory effect on HIVΔvif and the MLVs that were tested. The antiretroviral activity of mA3 splicing isoforms clearly impaired the infectious cycle of the AKV-N, AKV-NB, and MoMLV gammaretroviruses during the early stages of the infection. Expression levels of the mA3 splicing isoforms have also been shown in the producer cells (Figure-20,-21, -22, -23). As a result, the current experimental data support the conclusion that the mA3 protein is a potent restriction factor that acts against the replicative murine gammaretroviruses and single-cycle pseudoviruses like HIVΔvif and M5P MLV. In addition, these data support the general conception that the endogenous mA3 protein in mice plays a crucial role in preventing the transmission of retroviral pathogens from one animal to the other. Experiments with HIVΔvif also revealed that the mA3 protein is capable of inhibiting retroviruses of other species such as humans (Figure-25).

Furthermore, previous reports on murine gammaretroviruses revealed controversial results on the restriction of MLVs by mA3. I closely looked into these issues by attempting to identify the reasons behind the conflicting data in the literature. While many studies could not detect any antiviral activity of the mA3 protein against the MLVs, other reports showed that there is a considerable amount of restriction, but oddly without any detectable mutation in the viral genome (2, 6, 14, 25, 33, 49, 58). Unless there are other unknown restriction mechanisms of A3 proteins, my explanation for this antiretroviral activity was the potential deamination-independent restriction mechanisms of the mA3 protein. Previous studies have shown that

A3 proteins may disrupt the activity of tRNA primer binding, reverse transcriptase and other mechanisms in the reverse transcription and integration microenvironment, therefore, any of these mechanisms can be a potential target for deamination-independent restriction pathway of A3 proteins and cause substantial reductions on the infectivity levels at the early stages of retroviral infection (6, 7, 22, 26).

In my experiments, the mA3 protein inhibited the infectivity of MoMLV without G-to-A transition mutations in the viral genome (Figure-22 and Table-8). When the infectivity levels of both MoMLV and M5P MLV were compared, each of these viruses showed different sensitivities to mA3-mediated restriction (Figure-22, -23). Unless higher quantities of mA3 were transfected in the producer cells, M5P MLV always escaped from the mA3-mediated restriction (Figure-23, -24). MoMLV is a replicative murine gammaretrovirus that closely resembles the natural virus, whereas the M5P MLV is a single-cycle VSVG-pseudotyped pseudovirus produced only by multi-plasmid transfections. One important difference between these two viruses is that MoMLV genomic RNA is transcribed from its 5' LTR promoter and expresses the MLV *Env* gene. The genomic RNA of M5P MLV, and that of the gag-pol polyprotein and the VSVG glycoprotein, are all expressed from CMV promoters on their respective plasmids. As a result, M5P MLV produces an abundance of viral proteins that are disproportionate to those produced by MoMLV. These important differences between these two viruses, although reasonable, this conjecture may explain why more mA3 is needed to restrict M5P MLV as compared to MoMLV.

I also used AKV-NB in our experiments, which is a slightly different gammaretrovirus compared to MoMLV, to confirm and validate our findings. Even though MoMLV and AKV-NB share more than 80% sequence identity, these two viruses are differentially

sensitive to the deamination-dependent restriction mechanism of mA3 (35). As a result of this observation, the third objective of the project was designed to investigate whether the glycosylated form of the gag polyprotein has a role in protecting the viruses from the mA3-mediated enzymatic activity. The results from the viral mutation analysis indicated that the integrated genome of AKV-NB into NIH 3T3 cells contained mutations whereas the MoMLV and hybrid gammaretrovirus sequences had no G-to-A editing in their DNA (Table-8). It remains a mystery which part of the MoMLV protects the virus from the enzymatic function of the mA3 protein, but our results clearly show that the glycosylation and the glycogag protein of the gammaretroviruses do not hinder the deamination-dependent restriction pathway of the protein. This however does not exclude the possibility that glycosylation of gag may enhance the infectivity of these retroviruses *in vivo* through alternate mechanisms (Figure-42, -43, and Table-8).

Although AKV-NB was the only gammaretrovirus found to have detectable mutations in its genome, other MoMLV-based viruses were also strongly inhibited by the deamination-independent restriction mechanisms during the early stage of the infection. My results indicate that although the enzymatic activity of mA3 is weaker than that of hA3G, this protein has evolved additional effective mechanisms to combat retroviral pathogenicity. In conclusion, I identified important differences in the restriction sensitivity of various gammaretroviruses to mA3 and have observed the effective but transient early intracellular block against infection by mA3 by using our *in vitro* experimental model.

In addition, the antiretroviral activity of the mA3 was dissected by examining the structural organization as well as the function of each of the catalytic domains of the protein. I then took our work a step further by looking at the crucial residues within the mA3 catalytic

domains and their effect on restriction. The function of each of catalytic domain was investigated using infectivity assays where mA3 was mutated in one or both of its domains (Figure-26, -27, -28, -29, -30, -31). Experiments revealed that the inactivation of either the N-Terminal or C-Terminal catalytic domain of the protein affected the antiviral activity of the protein, although only slight changes were observed with replicative gammaretroviruses.(Figure-26, -27, -28, -29, -30, -31). Therefore, M5P MLV and HIVΔvif pseudoviruses were very useful during the infectivity assays with mA3 point mutants by providing an accessible platform to analyze the antiviral activity of these mutants clearly. This shows that both domains are involved in the restriction mechanisms of the protein. The deamination-independent restriction was preeminent over the deaminase-dependent activity due to the fact that inactivation of the catalytic activity of the protein did not have a drastic effect on its antiviral activity (Figure-26, -27, -28, -29, -30, -31). As a result, I decided that viral mutation analysis was required to better understand the involvement of mA3 catalytic domain mutations and their effect on viral restriction.

Determining the extent of DNA editing activity by mA3 on viral DNA was essential for the importance of the deamination-dependent restriction pathway of the mA3 protein splicing isoforms. My HIVΔvif laboratory strain was selected for these experiments since the targeted eGFP reporter gene is located at the ideal position to detect mutations in the viral genome (78) (Table-3). Analysis of the deaminase activity of mA3 indicated that the domain organization of the protein is reversed compared to hA3G. However, this structural diversity did not affect the enzymatic activity of the protein on HIVΔvif because the mutation rate of the mA3 splicing isoforms, except the low expresser mA3[Balb], were the same as hA3G (Table-3, -4). Interestingly, even though the antiretroviral activity of the mA3[C57]ΔE5-

E73A (first domain) and mA3[C57]ΔE5-E257A (second domain) point mutants were similar, there were no detectable mutations with the mA3[C57]ΔE5-E73A mutant showing that the second domain of the protein has no enzymatic activity but is still as effective (Figure-31). Also, mutation analysis indicated that the enzymatic activity of the mA3[C57]ΔE5-E257A mutant was 50% less than the wild-type protein (Table-4). As a result, I speculated that the second domain of the protein may have a role in facilitating the enzymatic activity of the first domain or is perhaps involved in the deamination-independent restriction mechanism.

Surprisingly, during the infectivity assays with HIVΔvif, I discovered that only the mA3 splicing isoforms without exon 5 were able to inhibit the infectivity of the virus and that the full-length mA3 protein splicing isoforms, mA3[C57] and mA3[Balb], only induced minor reductions in the infectivity levels of the virus (Figure-25). Analysis on the enzymatic activity of the mA3 splicing isoforms revealed that the full-length mA3 also induced a high mutation rate against HIVΔvif (Table-3). However, the potent DNA-editing activity of full-length mA3[C57] did not correlate with the restriction of the infectivity levels of the virus (Figure-25). The other full-length splicing isoform, mA3[Balb], also showed a relatively high mutation rate within HIVΔvif viral sequences, but the mutation rate was lower than the other splicing isoforms (Table-3). The inhibitory effects of this protein was also as low as the full-length mA3[C57] protein in my experiments (Figure-25).

My infectivity assays on various retroviruses and subsequent mutation analysis on viral sequences indicated that mA3 protein has a more complex restriction mechanism than previously described in the literature. It is reasonable to speculate that different restriction mechanisms of the mA3 work in conjunction with each other, in that, each mechanism has a

main antiretroviral function but can also facilitate the activity of the other mechanism. Previous reports indicated that A3 proteins cause random G-to-A mutations in the retroviral sequences and impair the infectivity of the virus (78). However, my results indicate that DNA editing does not always correlate with the antiretroviral activity against the virus and that heavily mutated viruses are still able to continue their life cycle (Figure-25, Table-3). Why exon 5 of mA3 disrupts the antiretroviral activity of the protein is currently unknown, but it is clear that the existence of this exon only affects the deamination-independent restriction mechanisms since the anti-HIVΔvif activity of the full-length mA3 splicing isoforms is almost equal to the other splicing isoforms. Also, the low antiviral activity of the full-length mA3 splicing isoforms was observed only against HIVΔvif. These full-length splicing isoforms were as effective as their shorter counterparts against MoMLV and AKV-NB. As a result, it is still unknown whether the heavy mutation load in viral genome or a single mutation in a critical coding sequence causes more damage on the infectivity of the virus.

After the initial discovery of the deamination-independent restriction pathway of the mA3, a catalytically inactive hA3G protein (hA3G-E69A/E259A) was generated and allowed us to examine whether the deamination-independent mechanism can also be observed in the human counterpart. My experiments with the hA3G-E69A/E259A protein and five different viruses indicated that the viral restriction activity of the hA3G-E69A/E259A protein remains partially active when the catalytic domains are inactivated (Figure-26, -27, -28, -29, -30). Expression levels of the hA3G-E69A/E259A protein were also similar to that of the wild type hA3G protein (Figure-32). I conclude that the deamination-independent restriction mechanism is likely conserved among A3 family members from divergent mammalian

species and that further investigations are required to understand the mechanisms behind this elusive phenomenon.

The mA3 splicing isoforms and the mA3[C57]ΔE5 point mutants were also analyzed for cytoplasmic complex formation. I was curious as to whether or not the splicing isoforms and the point mutations disrupt high molecular mass complex formations seen with the wild-type protein. My results show that these modifications did not hinder complex formation, or the overall expression of these proteins. In conclusion, all of the mA3 protein samples formed the HMM complexes in cell lines *in vitro* (Figure-33, -34).

For many years, it was thought that the enzymatic activity of the A3 protein was the only pathway leading to retroviral restriction. The identification of the strong deamination-independent restriction mechanism of the mA3 protein was a very important part of this study because it helped to dissect the antiviral functions of the protein and indicated that there are other functions of the protein that remain to be discovered. Previous reports on the hA3G protein also referred to the deamination-independent restriction pathway, however much controversy remains over this issue, with proponents claiming that enzymatic activity is the main component of the antiretroviral function of the protein (6, 7, 18, 22, 23, 26, 58, 67). Several studies looked into this non-enzymatic activity, and have concluded that hA3G interferes with reverse transcription and integration of the viruses (6, 7, 22, 26, 49). In addition, it was also proposed that interaction of A3 with the tRNA primer binding site may also interrupt viral replication and decrease infectivity levels (22). It is not very clear when or how the A3 proteins disrupt the replication cycle of retroviral pathogens, and the deamination-independent restriction mechanism is still unknown.

The catalytically inactive double domain mutant of hA3G (hA3G-E69A/E259A) and mA3[C57]ΔE5 (mA3[C57]ΔE5-E73A/E257A) were generated to distinguish the potency of the deamination-independent restriction mechanisms. Both of these mutants were tested on AKV-NB (Figure-35), and on HIVΔvif (Figure-36). I found that both catalytically inactive mutants of mA3 (mA3[C57]ΔE5-E73A/E257A) and hA3G (hA3G-E69A/E259A) provided strong deamination-independent restriction against AKV-NB (Figure-35) but this activity was weaker against HIVΔvif (Figure-36). It is now clear that both mA3 and hA3G proteins evolved in a way that allowed them to inhibit the retroviral replication cycle using more than one mechanism and that their activity is not solely dependent on deamination.

In order to investigate the intensity of the deamination-independent restriction pathway, replicative viruses, AKV-NB and MoMLV, were tested against both wild-type and catalytically inactive double domain mutants of mA3 (mA3[C57]ΔE5-E73A/E257A) and hA3G (hA3G-E69A/E259A) in a time course experiment to analyze A3-mediated restriction during the late stages of infection (up to 72 hours post-infection). I found that the deamination-independent restriction pathway of mA3 and hA3G was only effective at the early stages of infection (24 hours) (Figure-37) and the replicative gammaretroviruses were able to overcome the inhibitory effect of the deamination-independent mechanism and recovered high infectivity levels at later stages post- infection (72 hours) (Figure-38, -39). In contrast, the restriction mediated by the wild-type mA3 and hA3G was definitive and the replicative viruses could not recover from the antiretroviral activity of these proteins even after 72 hours (Figure-37, -38, -39). In conclusion, the results of these experiments on the deamination-dependent and –independent restriction mechanisms of the mA3 and hA3G

proteins indicated that both work together in order to inhibit the infectious cycle of the retroviral pathogens.

Current experimental data clearly supports that the deamination-independent restriction mechanism of mA3 and hA3G is a component of the antiretroviral activity of these proteins. In addition, the enzymatic activity of the A3 proteins is required for the long-term inhibition of retroviral pathogens. While A3 proteins provide an early intracellular block to retroviruses, the long term effects of the restriction may play less of a role since there are also other factors in the immune system that are involved in blocking the infectivity of these pathogens. However, the antiretroviral activity of the A3 proteins is very important upon initial invasion of retroviruses and subsequent cell modification events occur. The restrictive role of A3 proteins has drastic effects on prevention of retrovirus integration and therefore subsequent genome modification events, after which our adaptive immune system is no longer capable of inhibiting.

Lastly, DNA target sequence specificity of the mA3 splicing isoforms and the hA3G protein was investigated on a variety of retroviral sequences. A3 proteins act on cytidines in a specific DNA context. Each mutation in my viral DNA sequences was analyzed for mA3 or hA3G DNA context specificity and the results indicated that the mutations indeed had the same pattern expected from A3 catalytic activity. This confirmed that these were not random or spontaneous deamination of cytidines (Table-5, -6, -7, -9, -10).

In conclusion, my analysis of the mA3 protein revealed that it provides a potent early block to infection by human and mouse retroviruses and that this restriction is provided by two separate but complementary mechanisms: a deamination-dependent and -independent

restriction. My results also indicate that mutations introduced in retroviral DNA replication intermediates may not be the direct cause of the restriction of these pathogens. Other cellular factors or proteins may be involved in recognizing the uracils introduced by the enzymatic activity of mA3 and promoting the degradation of retroviral replication intermediates.

7.0 CONCLUSIONS

This study demonstrated that mA3 has a strong activity against retroviral pathogens. Analysis of the structural organization of the mA3 protein revealed differences when compared the human counterparts, whereas the antiviral activity yielded similar results. My investigations showed that allelic differences and splicing isoforms, of the mA3 protein have the most impact on the restrictive capabilities of the protein. Previous reports highlighted the importance of allelic variation among the mA3 proteins, but here I show rather that splicing isoforms of the protein are what have the most important impact on the antiretroviral activity. Also, the absence of exon 5 seems to improve the antiretroviral function of the protein.

In my study, expression level analysis of the mA3 splicing isoforms and careful titration of the retroviruses helped me to design an experimental model in which I could observe the inhibitory effects of the mA3 proteins after a single infection cycle. This model was very useful in identifying the restriction capabilities of different splicing isoforms and their activity against a variety of retroviruses. Experiments on human and mouse retroviruses also showed that mA3 is active against both human and mouse retroviruses, but has a stronger enzymatic activity against the HIV-1 virus than mouse gammaretroviruses. As a result, mA3 in mice can be a barrier for mouse retroviruses, but has also the potential to block zoonotic transmission of retroviral pathogens from other species as well.

My investigations of deamination-dependent and -independent restriction pathways of mA3 helped me to characterize the long-term restriction of retroviruses by APOBEC3 proteins. My results indicate that both the deamination-dependent and -independent mechanisms have an additive effect on the inhibition of the replication cycle of susceptible retroviral

pathogens. Together, these mechanisms provide an almost complete and durable block to the infectivity of the retroviruses. When both catalytic domains of human A3G and mA3[C57]ΔE5 were inactivated, antiretroviral activity still remained in my experiments, albeit a short-term activity that appeared to slow down infection kinetics rather than permanently damage the virus. This indicated that both enzymatic and non-enzymatic restriction mechanisms of A3 proteins are necessary for the complete inhibition of the retroviral infectious cycle.

Furthermore, experiments on the glycogag segment of the gammaretroviruses did not reveal any differences between the wild-type MoMLV virus and its hybrids with AKV-NB in regards to the deamination of viral sequences. This glycogag segment may provide advantages during the release of the viral particles, but does not shield such viruses from the enzymatic activity of mA3.

The results of this study demonstrate that mA3 has a variety of restriction mechanisms to impair the replication cycle of retroviral pathogens, and that this protein provides a solid intracellular block for the transmission and endogenization of retroviral pathogens. In my study, characterization of the structural organization of mA3 and identification of distinct restriction mechanisms of the protein greatly improved our current knowledge of this intrinsic restriction factor. The outcome of this study has already changed our ideas on how these proteins work against retroviral pathogens. In the future, my results will provide guidelines for further *in vivo* investigations on A3 knock-out and transgenic mouse models aiming at testing the relevance of these restriction factors under physiological conditions, as we have yet to understand how A3 proteins are regulated *in vivo*. Future work will involve designing a mouse model to study the restriction mechanisms of the A3 proteins *in vivo*, and

also, try to better understand the interactions of HIV-1 Vif and hA3G *in vivo* for the development of novel antiviral drugs against the virus.

8.0 REFERENCES

1. **Aagaard, L., Mikkelsen, J.G., Warming, S., Duch, M., and F.S. Pedersen.** 2002. Fv1-like restriction of N-tropic replication-competent murine leukaemia viruses in mCAT-1-expressing human cells. Journal of General Virology **83**:439-442.
2. **Abudu, A., Takaori-Kondo, A., Izumi, T., Shirakawa, K., Kobayashi, M., Sasada, A., Fukunaga, K., and T. Uchiyama.** 2006. Murine retrovirus escapes from murine APOBEC3 via two distinct novel mechanisms. Current Biology **16**:1565-1570.
3. **Albin, J.S., and R.S. Harris.** 2010. Interactions of host APOBEC3 restriction factors with HIV-1 in vivo: implications for therapeutics. Expert Reviews in Molecular Medicine **12**:e4.
4. **Barde, I., Salmon, P., and D. Trono.** 2010. Production and titration of lentiviral vectors. Current Protocols in Neuroscience **53**:4.21.1-4.21.23.
5. **Beale, R.C.L.** 2006. DNA sequence specificity of APOBEC family deaminases. Trinity College, Cambridge University.
6. **Bishop, K.N., Holmes, R.K., and M.H. Malim.** 2006. Antiviral potency of APOBEC proteins does not correlate with cytidine deamination. Journal of Virology **80**(17):8450-8458.
7. **Bishop, K.N., Verma, M., Kim, E-Y., Wolinsky, S.M., and M.H. Malim.** 2008. APOBEC3G inhibits elongation of HIV-1 reverse transcripts. PLoS Pathogens **4**(12): e1000231.
8. **Browne, E.P., and D.R. Littman.** 2008. Species-specific restriction of APOBEC3-mediated hypermutation. Journal of Virology **82**(3):1305-1313.
9. **C.A. Kozak.** 2010. The mouse "xenotropic" gammaretroviruses and their XPR1 receptor. Retrovirology **7**:101.
10. **Chen, B.K., Rousso, I., Shim, S., and P.S. Kim.** 2001. Efficient assembly of an HIV-1/MLV gag-chimeric virus in murine cells. Proceedings of the National Academy of Sciences of the United States of America **98**(26):15239-15244.
11. **Chiu, Y-L., and W.C. Greene.** 2006. Multifaceted antiviral actions of APOBEC3 cytidine deaminases. Trends in Immunology **27**(6):291-297.
12. **Chiu, Y-L., and W.C. Greene.** 2008. The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. The Annual Review of Immunology **26**:317-353.
13. **Coffin, J.M., Hughes, S.H., and H.E. Varmus.** 2002. Retroviruses.
14. **Conticello, S.G., Thomas, C.J.F., Petersen-Mahrt, S.K., and M.S. Neuberger.** 2005. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. Molecular Biology and Evolution **22**(2):367-377.
15. **Conticello, S.G., Langlois, M-A., and M.S. Neuberger.** 2007. Insights into DNA deaminases. Nature Structural & Molecular Biology **14**(1):7-9.
16. **Cordaux, R., and M.A. Batzer.** 2009. The impact of retrotransposons on human genome evolution. Nature Reviews Genetics **10**:691-703.

17. **Dewannieux, M., Ribet, D., and T. Heidmann.** 2010. Risks linked to endogenous retroviruses for vaccine production: A general overview. Biologicals **38**:366-370.
18. **Doehle, B.P., Schäfer, A., Wiegand, H.L., Bogerd, H.P., and B.R. Cullen.** 2005. Differential sensitivity of murine leukemia virus to APOBEC3-mediated inhibition is governed by virion exclusion. Journal of Virology **79**(13):8201-8207.
19. **Feng, Y., and L. Chelico.** 2011. Intensity of deoxycytidine deamination of HIV-1 proviral DNA by the retroviral restriction factor APOBEC3G is mediated by the non-catalytic domain. The Journal of Biological Chemistry **286**(13): 11415-11426.
20. **G.L. Buchschacher, Jr.** 2001. Introduction to retroviruses and retroviral vectors. Somatic Cell and Molecular Genetics **26**(1-6):1-11.
21. **Groom, H.C.T., Yap, M.W., Galâo, R.P., Neil, S.J.D., and K.N. Bishop.** 2010. Susceptibility of xenotropic murine leukemia virus-related virus (XMRV) to retroviral restriction factors. Proceedings of the National Academy of Sciences of the United States of America **107**(11):5166-5171.
22. **Guo, F., Cen, S., Niu, M., Saadatmand, J., and L. Kleiman.** 2006. Inhibition of $tRNA_{lys/3-}$ primed reverse transcription by human APOBEC3G during human immunodeficiency virus type 1 replication. Journal of Virology **80**(23):11710-11722.
23. **Hakata, Y., and N.R. Landau.** 2006. Reversed functional organization of mouse and human APOBEC3 cytidine deaminase domains. The Journal of Biological Chemistry **281**(48):36624-36631.
24. **Harris, R.S., and M.T. Liddament.** 2004. Retroviral restriction by APOBEC proteins. Nature Reviews Immunology **4**: 868-877.
25. **Hatziioannou, T., Cowan, S., and P.D. Bieniasz.** 2004. Capsid-dependent and –independent postentry restriction of primate lentivirus tropism in rodent cells. Journal of Virology **78**(2):1006-1011.
26. **Holmes, R.K., Malim, M.H., and K.N. Bishop.** 2007. APOBEC-mediated viral restriction: not simply editing? TRENDS in Biochemical Sciences **32**(3):118-128.
27. **Hye, B.E., Park, S-H., Park, S.M., Park, J.W., Lim, M.S., and Y-T. Jung.** 2008. Generation and characterization of a stable full-length ecotropic murine leukemia virus molecular clone that produces novel phenotypes to Fv1 restriction. Journal of Microbiology and Biotechnology **18**(4):799-804.
28. **Jern, P., Stoye, J.P., and J.M. Coffin.** 2007. Role of APOBEC3 in genetic diversity among endogenous murine leukemia viruses. PLoS Pathogens **3**(10):e183.
29. **Johal, H., Faedo, M., Faltas, J., Lau, A., Mousina, R., Cozzi, P., deFazio, A., and W.D. Rawlinson.** 2010. DNA of mouse mammary tumor virus-like virus is present in human tumors influenced by hormones. Journal of Medical Virology **82**:1044-1050.
30. **Kirchhoff, F.** 2010. Immune evasion and counteraction of restriction factors by HIV-1 and other primate lentiviruses. Cell Host & Microbe **8**(1):55-67.
31. **Knoper, R.C., Ferrarone, J., Yan, Y., Lafont, B.A.P., and C.A. Kozak.** 2009. Removal of either N-glycan site from the envelope receptor binding domain of Moloney and Friend but not AKV mouse ecotropic gammaretroviruses alters receptor usage. Virology **391**:232-239.

32. **Kolokithas, A., Rosenke, K., Malik, F., Hendrick, D., Swanson, L., Santiago, M.L., Portis, J.L., Hasenkrug, K.J., and L.H. Evans.** 2010. The glycosylated gag protein of murine leukemia viruses inhibits the anti-retroviral function of APOBEC3. Journal of Virology **84**(20):10933-10936.
33. **Krishna, D., and J.M. Le Doux.** 2006. Murine leukemia virus particles activate Rac1 in HeLa cells. Biochemical and Biophysical Research Communications **345**:1184-1193.
34. **Langlois, M-A., and M.S. Neuberger.** 2008. Human APOBEC3G can restrict retroviral infection in avian cells and acts independently of both UNG and SMUG1. Journal of Virology **82**(9):4660-4664.
35. **Langlois, M-A., Kemmerich, K., Rada, C., and M.S. Neuberger.** 2009. The AKV murine leukemia virus is restricted and hypermutated by mouse APOBEC3. Journal of Virology **83**(22):11550-11559.
36. **Langlois, M-A., Beale, R.C.L., Conticello, S.G., 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 Research **33**(6):1913-1923.
37. **LaRue, R.S., Jónsson, S.R., Silverstein, K.A.T., Lajoie, M., Bertrand, D., El-Mabrouk, N., Hòtzel, I., Andrésdóttir, V., Smith, T.P.L., and R.S. Harris.** 2008. The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. BMC Molecular Biology **9**:104.
38. **Lever, A.M.L., and K-T. Jeang.** 2011. Insights into cellular factors that regulate HIV-1 replication in human cells. Biochemistry **50**: 920-931.
39. **Lo, S-C., Pripuzova, N., Li, B., Komaroff, A.L., Hung, G-C., Wang, R., and H.J. Alter.** 2010. Detection of MLV-related virus gene sequences in blood of patients with chronic fatigue syndrome and healthy blood donors. Proceedings of the National Academy of Sciences of the United States of America **107**(36):15874-15879.
40. **Lombardi, V.C., Ruscetti, F.W., Gupta, J.D., Pfost, M.A., Hagen, K.S., Peterson, D.L., Ruscetti, S.K., Bagni, R.K., Petrow-Sadowski, C., Gold, B., Dean, M., Silverman, R.H., and J.A. Mikovits.** 2009. Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome. Science **326**:585-589.
41. **Low, A., Datta, S., Kuznetsov, Y., Jahid, S., Kothari, N., McPherson, A., and H. Fan.** 2007. Mutation in the glycosylated gag protein of murine leukemia virus results in reduced in vivo infectivity and a novel defect in viral budding or release. Journal of Virology **81**(8):3685-3692.
42. **Low, A., Okeoma, C.M., Lovsin, N., Heras, M., Taylor, T.H., Peterlin, B.M., Ross, S.R., and H. Fan.** 2009. Enhanced replication and pathogenesis of Moloney murine leukemia virus in mice defective in the murine APOBEC3 gene. Virology **385**:455-463.
43. **M-A. Langlois.** 2010. Mother's milk and intrinsic immunity. Cell Host & Microbe **8**(6):467-469.

44. **M.A. Sommerfelt.** 1999. Retrovirus receptors. Journal of General Virology **80**:3049-3064.
45. **M.L. Santiago.** 2008. APOBEC3 encodes Rfv3, a gene influencing neutralizing antibody control of retrovirus infection. Science **321**:1343-1346.
46. **M. Pizzato.** 2010. MLV glycosylated-gag is an infectivity factor that rescues nef-deficient HIV-1. Proceedings of the National Academy of Sciences of the United States of America **107**(20):9364-9369.
47. **Ma, J., Li, X., Xu, J., Zhang, Q., Liu, Z., Jia, P., Zhou, J., Guo, F., You, X., Yu, L., Zhao, L., Jiang, J., and S. Cen.** 2011. The cellular source for APOBEC3G's incorporation into HIV-1. Retrovirology **8**:2.
48. **Malim, M.H., and M. Emerman.** 2008. HIV-1 accessory proteins – ensuring viral survival in a hostile environment. Cell Host & Microbe **3**:388-398.
49. **Mbisa, J.L., Barr, R., Thomas, J.A., Vandegraaff, N., Dorweiler, I.J., Svarovskaia, E.S., Brown, W.L., Mansky, L.M., Gorelick, R.J., Harris, R.S., Engelman, A., 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. Journal of Virology **81**(13): 7099-7110.
50. **Menéndez-Arias, L.** 2011. Evidence and controversies on the role of XMRV in prostate cancer and chronic fatigue syndrome. Reviews in Medical Virology **21**:3-17.
51. **Mikl, M.C.** 2005. Genetic analysis of APOBEC2 and APOBEC3 in mice. MRC Laboratory of Molecular Biology & Trinity College, Cambridge University.
52. **Mikl, M.C., Watt, I.N., Lu, M., Reik, W., Davies, S.L., Neuberger, M.S., and C. Rada.** 2005. Mice deficient in APOBEC2 and APOBEC3. Journal of Virology **25**(16):7270-7277.
53. **Miller, A.D., and M.J. Metzger.** 2011. APOBEC3-mediated hypermutation of retroviral vectors produced from some retrovirus packaging cell lines. Gene Therapy **18**(5): 528-530.
54. **Navaratnam, N., and R. Sarvar**. 2006. An overview of cytidine deaminases. International Journal of Hematology **83**:195-200.
55. **Nitta, T., Kuznetsov, Y., McPherson, A., and H. Fan.** 2010. Murine leukemia virus glycosylated gag (gPr80gag) facilitates interferon-sensitive virus release through lipid rafts. Proceedings of the National Academy of Sciences of the United States of America **107**(3):1190-1195.
56. **Okeoma, C.M., Huegel, A.L., Lingappa, J., Feldman, M.D., and S.R. Ross.** 2010. APOBEC3 proteins expressed in mammary epithelial cells are packaged into retroviruses and can restrict transmission of milk-borne virions. Cell Host & Microbe **8**(6):534-543.
57. **Okeoma, C.M., Low, A., Bailis, W., Fan, H.Y., Peterlin, B.M., and S.R. Ross.** 2009. Induction of APOBEC3 in vivo causes increased restriction of retrovirus infection. Journal of Virology **83**(8):3486-3495.
58. **Okeoma, C.M., Lovsin, N., Peterlin, B.M., and S.R. Ross.** 2007. APOBEC3 inhibits mouse mammary tumour virus replication in vivo. Nature **445**:927-930.

59. **Okeoma, C.M., Petersen, J., and S.R. Ross.** 2009. Expression of murine APOBEC3 alleles in different mouse strains and their effect on mouse mammary tumor virus infection. Journal of Virology **83**(7):3029-3038.
60. **Paprotka, T., Delviks-Frankenberry, K.A., Cingòz, O., Martinez, A., Kung, H-S., Tepper, C.G., Hu, W-S., Fivash Jr., M.J., Coffin, J.M., and V.K. Pathak.** 2011. Recombinant origin of the retrovirus XMRV. Science June 2nd 10.1126/science.1205292.
61. **Refsland, E.W., Stenglein, M.D., Shindo, K., Albin, J.S., Brown, W.L., and R.S. Harris.** 2010. Quantitative profiling of the full *APOBEC3* mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. Nucleic Acids Research **38**(13):4274-4284.
62. **Ross, S.R.** 2003. Mouse mammary tumor virus and the immune system. Immunologic Research **27**(2-3):469-479.
63. **Ross, S.R.** 2008. MMTV infectious cycle and the contribution of virus-encoded proteins to transformation of mammary tissue. Journal of Mammary Gland Biol Neoplasia **13**(3):299-307.
64. **Ross, S.R.** 2009. Are viruses inhibited by APOBEC3 molecules from their host species? PLoS Pathogens **5**(4):e1000347.
65. **Ross, S.R.** 2010. Mouse mammary tumor virus molecular biology and oncogenesis. Viruses **2**:2000-2012.
66. **Rowe, H.M., and D. Trono.** 2011. Dynamic control of endogenous retroviruses during development. Virology **411**:273-287.
67. **Rulli,Jr., S.J., Mirro, J., Hill, S.A., Lloyd, P., Gorelick, R.J., Coffin, J.M., Derse, D., and A. Rein.** 2008. Interactions of murine APOBEC3 and human APOBEC3G with murine leukemia viruses. Journal of Virology **82**(13):6566-6575.
68. **Rusmevichientong, A., and S.A. Chow.** 2010. Biology and pathophysiology of the new human retrovirus XMRV and its association with human disease. Immunology Research DOI 10.1007/s12026-010-8165-y.
69. **Santiago, M.L., Montano, M., Benitez, R., Messer, R.J., Yonemoto, W., Chesebro, B., Hasenkrug, K.J., and W.C. Greene.** 2008. APOBEC3 encodes rfv3, a gene influencing neutralizing antibody control of retrovirus infection. Science **321**: 1343-1346.
70. **Santoni de Sio, F.R., and D. Trono.** 2009. APOBEC3G-depleted resting $CD4^+$ T cells remain refractory to HIV-1 infection. PLoS One **4**(8):e6571.
71. **Stieler, K., and N. Fischer.** 2010. APOBEC3G efficiently reduces infectivity of the human exogenous gammaretrovirus XMRV. PLoS One **5**(7): e11738.
72. **Stieler, K., Schulz, C., Lavanya, M., Aepfelbacher, M., Stocking, C., and N. Fischer.** 2010. Host range and cellular tropism of the human exogenous gammaretrovirus XMRV. Virology **399**:23-30.
73. **Suspène, R., Henry, M., Guillot, S., Wain-Hobson, S., and J-P. Vartanian.** 2005. Recovery of APOBEC3-edited human immunodeficiency virus G→A hypermutants by differential DNA denaturation PCR. Journal of General Virology **86**:125-129.

74. **Takeda, E., Tsuji-Kawahara, S., Sakamoto, M., Langlois, M-A., Neuberger, M.S., Rada, C., and M. Miyazawa.** 2008. Mouse APOBEC3 restricts Friend leukemia virus infection and pathogenesis in vivo. Journal of Virology **82**(22):10998-11008.
75. **Urisman, A., Molinaro, R.J., Fischer, N., Plummer, S.J., Casey, G., Klein, E.A., Malathi, K., Magi-Galluzzi, C., Tubbs, R.R., Ganem, D., Silverman, R.H., and J.L. DeRisi.** 2006. Identification of a novel gammaretrovirus in prostate tumors of patients homozygous for R462Q *RNASEL* variant. PLoS Pathogens **2**(3):e25.
76. **Wang, X., Abudu, A., Son, S., Dang, Y., Venta, P.J., and Y-H. Zheng.** 2011. Journal of Virology **85**(7): 3142-3152.
77. **Yang, Z.** 2007. Assaying cytidine deamination. MRC Laboratory of Molecular Biology & Trinity College, University of Cambridge.
78. **Yu, Q., Kònig, R., Pillai, S., Chiles, K., Kearney, M., Palmer, S., Richman, D., Coffin, J.M., and N.R. Landau.** 2004. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. Nature Structural & Molecular Biology **11**(5):435-442.
79. **Zhang, L., Li, X., Ma, J., Yu, L., Jiang, J., and S. Cen.** 2008. The incorporation of APOBEC3 proteins into murine leukemia viruses. Virology **378**:69-78.

9.0 APPENDIX

Plasmid maps have been generated in order to describe the structure of the constructs. Vector NTI software (Invitrogen, Carlsbad, California, USA) was used to design the primers and expression vectors in the study.

Figure-37, -38, -39 (Line Chart for the AKV-NB virus): Time course infectivity comparison of the wild-type mA3[C57]ΔE5 and hA3G proteins with the catalytically inactive mutants of these two proteins against the AKV-NB virus. The A3 proteins were tested against the AKV-NB gammaretrovirus and the infectivity levels were measured by flow cytometry analysis. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population.

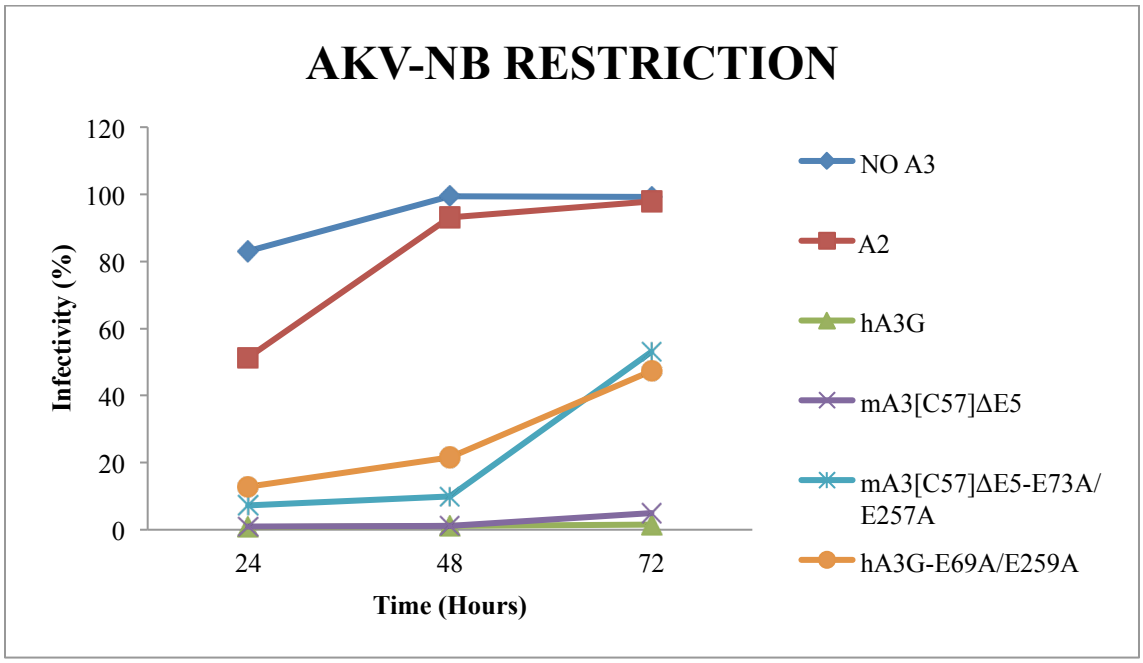

AKV-NB RESTRICTION
120
100
80
60
40
20
0
Infectivity (%)
24
48
72
Time (Hours)
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-E73A/
E257A
hA3G-E69A/E259A

Figure-37, -38, -39 (Line Chart for the MoMLV virus): Time course infectivity comparison of the wild-type mA3[C57]ΔE5 and hA3G proteins with the catalytically inactive mutants of these two proteins against the MoMLV virus. The A3 proteins were tested against the MoMLV gammaretrovirus and the infectivity levels were measured by flow cytometry analysis. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population.

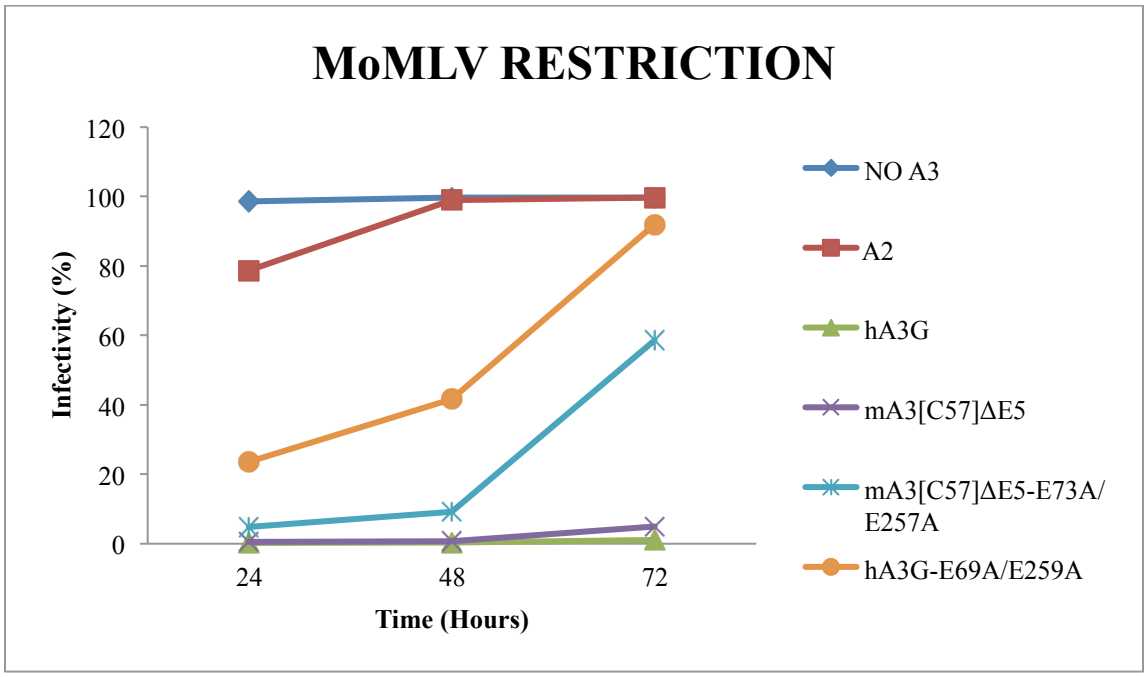
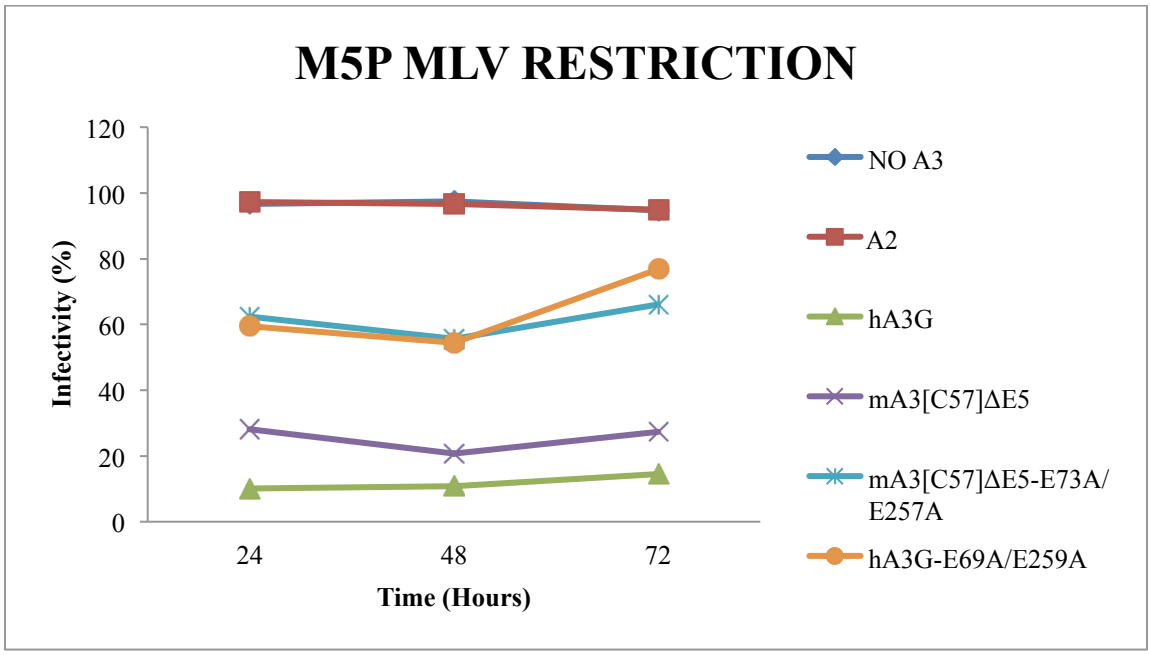

MoMLV RESTRICTION
120
100
80
60
40
20
0
Infectivity (%)
24
48
72
Time (Hours)
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-E73A/ E257A
hA3G-E69A/E259A

Figure-37, -38, -39 (Line Chart for the M5P MLV virus): Time course infectivity comparison of the wild-type mA3[C57]ΔE5 and hA3G proteins with the catalytically inactive mutants of these two proteins against the M5P MLV virus. The A3 proteins were tested against the M5P MLV gammaretrovirus and the infectivity levels were measured by flow cytometry analysis. Infected cells turn green as the result of eGFP expression from integrated retroviruses. Infectivity (%) is measured as the percentage of green cells in the analysed population.


M5P MLV RESTRICTION
120
100
80
60
40
20
0
Infectivity (%)
24
48
72
Time (Hours)
NO A3
A2
hA3G
mA3[C57]ΔE5
mA3[C57]ΔE5-E73A/E257A
hA3G-E69A/E259A