MOLECULAR MECHANISMS OF HUMAN INNATE IMMUNE RESTRICTION FACTORS
APOBEC3G AND APOBEC3H
FOR INHIBITING THE RETROVIRUS HIV-1 AND THE RETROTRANSPOSON LINE1

A Thesis Submitted to the College of
Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in the Department of Microbiology and Immunology
University of Saskatchewan
Saskatoon, Saskatchewan

By Yuqing Feng

PERMISSION TO USE

ABSTRACT

The APOBEC3 restriction factors belong to a family of cytidine deaminases that are able to suppress the replication of viruses with a single-stranded DNA intermediate by inducing mutagenesis and functional inactivation of the virus. Of the seven common human APOBEC3 enzymes, only APOBEC3-D, -F, -G and -H appear to be relevant to HIV-1 restriction in CD4$^+$ T cells. The restriction of HIV-1 by APOBEC3 enzymes occurs most potently in the absence of HIV-1 Vif, which counteracts APOBEC3s by inducing polyubiquitination and subsequent degradation of APOBEC3 enzymes. Virion-encapsidated APOBEC3s can deaminate cytosines to uracils in viral (-)DNA. Upon replication of (-)DNA to (+)DNA, the HIV-1 reverse transcriptase incorporates adenines opposite to the uracils hereby inducing C/G to T/A transition mutations. Among all APOBEC3 enzymes that are relevant to HIV-1 restriction, APOBEC3G is the most studied APOBEC3 enzyme. APOBEC3G has been shown to processively catalyze deamination reactions on single-stranded (-)DNA using a mechanism called facilitated diffusion, which involves sliding and jumping movements in search of target cytosine-containing motifs. This single-stranded DNA scanning mechanism allows APOBEC3G to efficiently deaminate multiple cytosines within one enzyme-DNA encounter and it is important for the mutational inactivation of HIV-1 *in vivo.* Vif attempts to neutralize APOBEC3G's function not only by inducing proteasomal degradation, but also by several degradation-independent mechanisms, such as inhibiting APOBEC3G virion encapsidation, mRNA translation, and for those APOBEC3G molecules that still become virion encapsidated, Vif has been shown to inhibit APOBEC3G's deamination activity.

My Ph.D. thesis work investigated the molecular mechanism of degradation-independent Vif-mediated inhibition of APOBEC3G and APOBEC3H deamination activity. This research led to the development of the hypothesis that Vif has developed a unique interaction with each APOBEC3 enzyme due to the different selection pressures they impose on HIV-1. Thus, we investigated how the interaction of Vif differs between APOBEC3G and APOBEC3H and characterized the activity of APOBEC3H as a restriction factor. This research allowed us to have a better understanding of the molecular determinants that govern an efficient APOBEC3-degradatation by HIV-1 Vif and provide insights for APOBEC3-based HIV-1 therapeutics.

Two Vif variants obtained from HIV-1 laboratory isolates, Vif$_{HXB2}$ and Vif$_{IIIB}$, were used to determine the degradation-independent effects of Vif on APOBEC3G. Biochemical assays

using a model HIV-1 replication assay and synthetic single-stranded or partially double-stranded DNA substrates demonstrated that APOBEC3G has an altered processive mechanism in the presence of Vif, and this caused APOBEC3G to be less effective at inducing mutagenesis in a model HIV-1 replication assay.

APOBEC3H is unique in that it is the only single-domain common APOBEC3 enzyme that restricts HIV-1. APOBEC3H exists in humans as seven haplotypes (I-VII) with different cellular stabilities. Amongst three stable APOBEC3H haplotypes, haplotype II and V occur most frequently in the population. I characterized the single-stranded DNA scanning mechanisms that haplotype II and V use to search their single-stranded substrate for cytosine-containing deamination motif. APOBEC3H haplotype II was able to processively deaminate its substrate using Brownian motion-driven movements termed sliding, jumping and intersegmental transfer, whereas haplotype V showed compromised sliding and intersegmental transfer abilities but was able to jump along single-stranded DNA. Since an Asp or Glu at amino acid 178 differentiates these two haplotypes, these data suggest this amino acid on predicted helix 6 contributes to processivity. I found the optimal processivity on ssDNA also required dimerization of APOBEC3H through the β2 strands. The diminished processivity of APOBEC3H haplotype V did not result in a reduced efficiency to restrict HIV-1 replication in single-cycle infectivity assay. This suggests a redundancy in the contribution of jumping and intersegmental transfer to mutagenic efficiency.

Vif$_{HXB2}$, but not Vif$_{IIIB}$, can cause degradation of APOBEC3H even though APOBEC3H interacts with both Vif variants. APOBEC3G is degraded after interaction with both of these Vif variants. To define the parameters for efficient Vif-induced degradation of an APOBEC3 enzyme, I used an *in vitro* quantitative method to determine the binding strength of APOBEC3G and APOBEC3H with Vif variant heterotetramers (Vif/CBFβ/EloB/EloC), the most stable form of Vif. Our biochemical analysis, along with cellular experiments to determine Vif-induced degradation efficiency and APOBEC3-Vif interactions in cells, support a model in which the degradation efficiency of Vifs correlates with both the APOBEC3-Vif binding strength and APOBEC3-Vif interface.

I also investigated how APOBEC3 enzymes restrict the replication of retrotransposon LINE-1. Retrotransposons are DNA sequences that replicate using a "copy-and-paste" mechanism through an RNA intermediate. The degradation of deaminated L1 cDNA rendered it

difficult to detect any APOBEC3-induced G-to-A mutations while the addition of uracil DNA glycosylase inhibitor allowed for the recovery of the APOBEC3-mediated deamination events. I found that two stable A3H haplotypes (haplotype II and haplotype V) use a deamination-independent mechanism to restrict L1 mobilization and compared the ability of APOBEC3H's to inhibit LINE1 with that of APOBEC3A and APOBEC3G, two APOBEC3s whose LINE1 restriction ability have been previously characterized. Taken together, these studies of the molecular mechanisms that APOBEC3G and APOBEC3H use to inhibit HIV-1 and LINE1 have allowed us to better understand their biological properties as cytidine deaminases and the determinants in APOBEC3s that made them efficient host innate immune restriction factors.

DEDICATION

I dedicate this thesis to my mentor,

Dr. Linda Chelico

Without whose support

this thesis would not have seen the light of the day

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

(+)DNA- Positive strand/coding strand

(-)DNA- Negative strand/template or non-coding strand

AID- Activation induced cytidine deaminase

AIDS- Acquired immunodeficiency syndrome

A3- APOBEC3

APOBEC3- Apolipoprotein B mRNA-editing enzyme-catalytic, polypeptide-like 3

APE- Apurinic/apyrimidinic endonuclease

ATP- Adenosine triphosphate

bp- Nucleotide base pair

C- Cytosine

CBFβ- Core binding factor β

cDNA- Complementary DNA

CTD- C terminal domain

CTL-Cytotoxic T lymphocyte

CRL5- Cullin-RING ubiquitin ligase 5

Cul5- Cullin5

dNTPs- Deoxyribonucleoside triphosphate

dsDNA- Double-stranded DNA

DTT- Dithiothreitol

E. coli- Escherichia coli

Env- Envolope

Elo B- Elongin B

Elo C- Elongin C

ERV- Endogenous retroviruses

FAM- Fluorescein

Hap 2- Haplotype 2

Hap5- Haplotype 5

HAART- Highly active anti-retroviral therapy

HIV-1- Human immunodeficiency virus type I

Ig- Immunoglobulin

IN- Integrase

K_d- Dissociation constant

LEAP- L1 element amplification protocol

LINE1- Long interspersed nuclear element

LTR- Long terminal repeat

mRNA- Messenger RNA

NC- Nucleocapsid

NTD- N terminal domain

nt- Nucleotides

ORF- Open reading frame

PAGE- Polyacrylamide gel electrophoresis

PBS- Primer binding site/phosphate-buffered saline

PBMC- Peripheral blood mononuclear cell
PPT- Polypurine tract
Prot- Protease
R- Nucleotide A or G
Rbx2- RING box protein 2
RNase H- Ribonuclease H domain enzyme
RNP- Ribonucleoprotein
RT- Reverse transcriptase
RUNX1- Runt-related transcription factor 1
rNTPs- Ribonucleotides
RT- Reverse transcriptase
SDM- Site directed mutagenesis
SDS-PAGE- Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SINE- Short interspersed nuclear element
SIVcpz- Simian immunodeficiency virus infecting chimpanzees
SNP- Single nucleotide polymorphism
SOCS2- Suppressor of cytokine signaling 2
ssDNA- Single-stranded DNA
TPRT- Target primed reverse transcription
UDG/UNG- Uracil DNA glycosylase
UGI- Uracil DNA glycosylase inhibitor
3'UTR-3' Untranslated regions
Vif-Viral infectivity factor
W- Nucleotides A or T
Z1- Zinc-coordinating domain type 1
Z2- Zinc-coordinating domain type 2
Z3- Zinc-coordinating domain type 3

1.0 Introduction and Literature Review
1.1 General HIV-1 biology
1.1.1 The basics of HIV replication

Human immunodeficiency virus type 1 (HIV-1, referred to as HIV) is a major global public health issue and currently infects approximately 37 million people worldwide. Sub-Saharan Africa is the most affected area and accounts for two-thirds of the global total of new HIV infections [1]. HIV targets immune cells, primarily CD4$^+$ T cells and macrophages, for infection and impairs the function of immune defence system against other infections. This can gradually render the infected individuals immunodeficient and eventually leads to acquired immunodeficiency syndrome (AIDS), the most advanced stage of HIV infection [2, 3]. Although highly active anti-retroviral therapy (HAART) can effectively control viraemia, variations of HIV that arise while a person is taking the medication can lead to drug-resistant strains of HIV and cause the therapy to fail [4]. Currently there is no cure for HIV infection.

HIV belongs to the genus *lentivirus*, part of the family *Retroviridae* [5]. Similar to all retroviruses, HIV contains two copies of non-covalently linked, positive-sense singled-stranded (ss) RNAs enclosed in a capsid composed of the viral protein p24 [6]. The HIV genome is 9.7-kb in length and it contains three genes that are essential for HIV replication: *gag* (group-specific antigen), *pol* (polymerase) and *env* (envelope) [7] (Fig. 1.1A). *Gag* encodes the precursor gag polyprotein which is subject to cleavage by the viral protease to generate the internal structural subunits of the virion including capsid, matrix and nucleocapsid; *pol* encodes the reverse transcriptase (RT), integrase, and protease enzymes that are crucial for genome replication, proviral integration, and gag polyprotein processing, respectively; *env* encodes the viral envelope protein which is essential for infection of target host cells [3]. HIV also encodes two regulatory proteins, the transcriptional transactivator (Tat) and the regulator of virion gene expression (Rev), which are required for transcriptional regulation and mRNA nuclear export, respectively. In addition, HIV contains four genes that encode accessory proteins: the viral protein r (Vpr), the viral protein u (Vpu), viral infectivity factor (Vif) and negative regulatory factor (Nef) [8-10]. HIV accessory proteins are not always required for viral replication, but instead are involved in enhancing viral fitness by allowing the viruses to evade host innate and adaptive immune mechanisms [8, 9]. These accessory proteins (Vpr, Vpu, Nef, and Vif) generally take advantage of the host protein degradative pathways to remove proteins such as restriction factors or

A

B

2
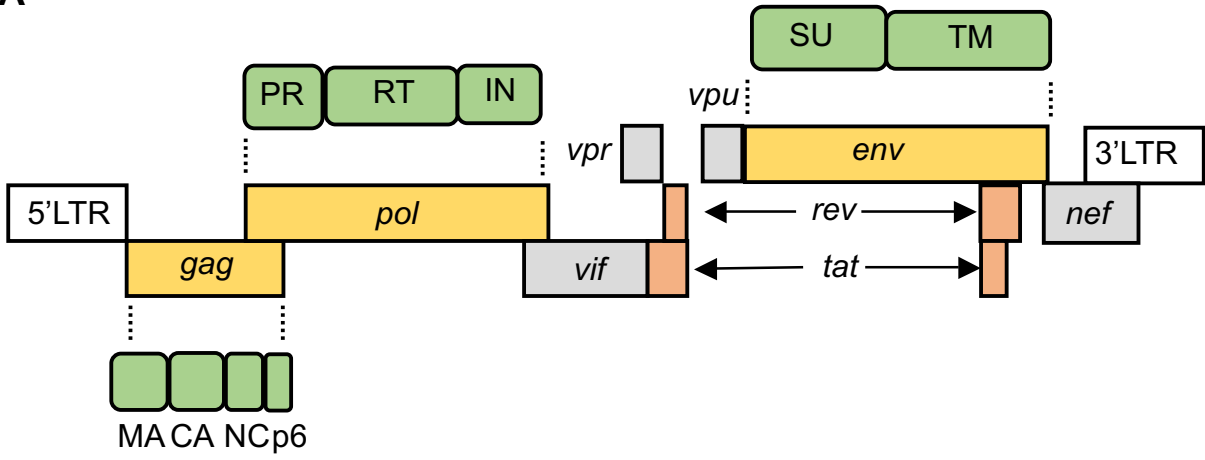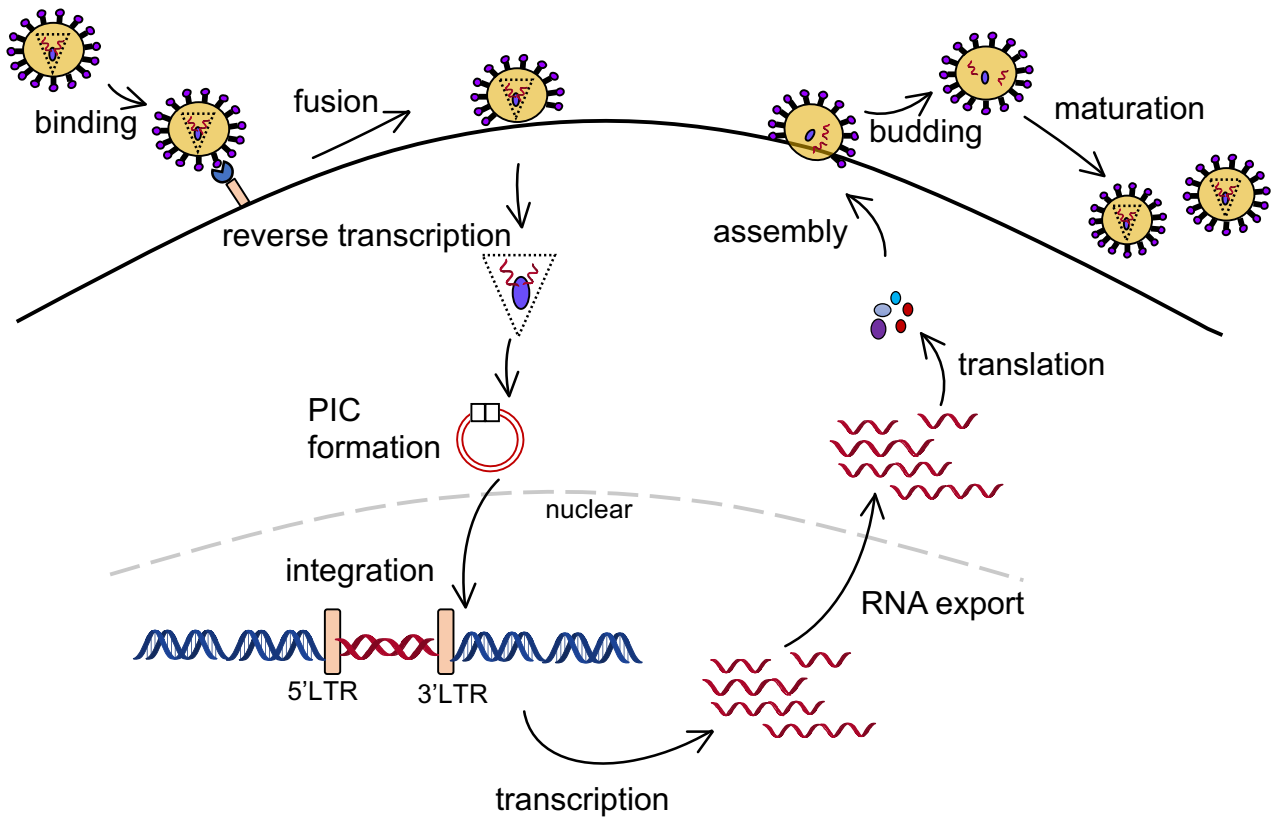
Figure 1.1 Schematic overview of the HIV genome structure and life cycle. (A) Schematic of HIV genome and encoded proteins. HIV contains three genes (yellow) that are common to all retroviruses. These three genes (*gag, pol and env*) encode polyproteins that will be subsequently processed to their mature forms (green). The 9.7-kb HIV genome contains open reading frames for 15 distinct proteins: the Gag and Env structural proteins MA (matrix), CA (capsid), NC (nucleocapsid), p6, SU (surface), and TM (transmembrane); the Pol proteins: PR (protease), RT (reverse transcriptase), and IN (integrase); two regulatory proteins Tat and Rev (orange); four accessory proteins Nef, Vif, Vpr, and Vpu (wheat) [11]. (B) HIV replication cycle. The figure illustrates the major steps involved in the HIV replication cycle. HIV envelope glycoprotein binds to the cell surface receptor, which triggers the fusion of the viral and cellular membranes and entry of the viral particle into the cell. Uncoating of the viral capsid, coupled with reverse transcription, leads to the formation of a double-stranded DNA and a pre-integration complex (PIC). PIC will be transported into the nucleus and with PIC-associated viral integrase (vertical box) forms an integrated provirus. Proviral DNA (red) transcription is mediated by the host RNA polymerase II, producing viral mRNA transcripts (red) of different sizes. These mRNAs serve as templates for viral protein production. After assembly, viral particle buds off and is released followed by protease-mediated maturation and results in the generation of infectious virions. The assembly virion contains the viral genome (two single-stranded RNA molecules), cellular tRNA$^{\text{lys,3}}$, some viral proteins such as Env, Gag polyprotein, NC, RT, IN, Vpr, and cellular proteins such as A3s [7, 12]. For simplicity, only one protein (shown in purple) is represented inside the virion structure.

immune signaling components that would otherwise interfere with the viral replication and propagation [8, 9]. For example, Vpr hijacks an E3 ubiquitin ligase complex and induces G2 cell cycle arrest, Vpu enhances viral particle budding by degradation and sequestration of the host restriction factor Bone marrow stromal cell antigen 2 (BST-2)/tetherin, and Nef downregulates h (MHC-1) cell surface expression to prevent viral antigen presentation and clearance of infected cells by cytotoxic T lymphocytes [8, 9]. Central to this thesis work is the accessory protein Vif. Vif mediates the degradation of the Apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3 (APOBEC3, referred to as A3) family of cytidine deaminases, which suppresses their ability to hypermutate and inactivate the viral genome. The function of Vif will be examined further in section 1.4.

The HIV replication cycle begins with the binding of the virus envelope glycoprotein gp120 to the CD4 receptor as well as to the chemokine co-receptor CC receptor type 5 (CCR5) or CXC receptor type 4 (CXCR4) on the surface of target CD4$^+$T cells, followed by fusion of the viral and cellular membranes and uncoating of capsid [13] (Fig. 1.1B). Although early research indicated that the capsid dissembles near the membrane immediately after the fusion [14, 15], a more favourable model proposes that capsid would remain intact until the incoming viral complexes reach the nuclear membrane and that uncoating occurs at the nuclear pore upon completion of reverse transcription, a process during which HIV converts its ss positive sense RNA genome into double-stranded DNA using host-encoded RT [13] (Fig 1.2). An intact capsid is important for keeping a high stoichiometry of HIV reverse transcriptase relative to the viral genome, thus enabling an efficient reverse transcription process. Because HIV capsid hexamers have a pore at the six-fold symmetry axis, they are permeable to small macromolecules such as deoxynucleoside triphosphates (dNTPs) and it is believed that HIV reverse transcription is triggered by the relatively high concentrations of dNTPs in the cytoplasm that diffuse through the inter-ring spacings of the capsid [16, 17]. Uncoating accompanies the formation of a HIV pre-integration complex (PIC), which is an integration-competent complex that is formed inside the cytoplasm upon viral dsDNA synthesis and can later integrate into the chromosome of the target cell and form the long terminal repeat (LTR)-flanked provirus with the assistance of the integrase enzyme [3, 13] (Fig. 1.1B). The 5'LTR (long terminal repeat) contains enhancer and promoter sequences which serves as the binding sites for host transcription factors that will later
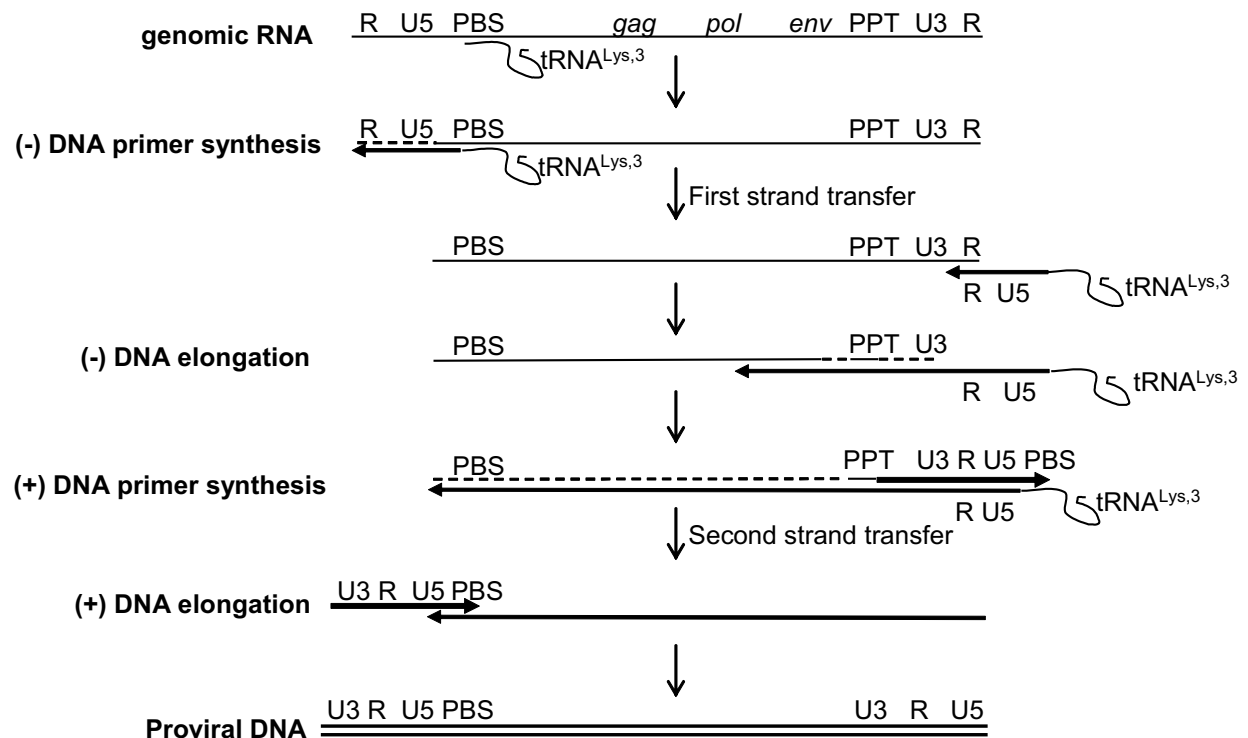
Figure 1.2 Schematic of reverse transcription and replication of HIV. The genomic RNA (black line) is primed for replication using a host tRNA^{lys3} primer. Synthesis of the (-)cDNA (bold line) generates an RNA/DNA hybrid that is a substrate for RNase H. RNase H degrades the genomic RNA strand (hatched line), leaving the nascent (-)cDNA single-stranded. Because both ends of the viral RNA genome contain the same repeated sequence (R), this allows the (-)cDNA to hybridize with the R sequence at the 3'end of the viral RNA, a step that is called the first strand transfer. With the hybridization of (-)cDNA to 3'R, (-)cDNA synthesis resumes along the viral RNA and RNase H degrades the RNA strand (hatched line) as the cDNA synthesis proceeds. A purine-rich sequence near the 3'end of the viral RNA genome called the polypurine tract or PPT is relatively resistant to RNase H degradation and serves as the primer for (+)cDNA synthesis. (+)cDNA synthesis proceeds until RT starts to copy the tRNA primer. With the completion of the 3' end tRNA reverse transcription, an RNA/DNA hybrid is created and RNase H removes the tRNA primer. This steps exposes a single-stranded portion of the (+)cDNA, which now has the same sequence as the 5' PBS. In a process called the second strand transfer, both the (+) and (-)cDNA are extended until the entire DNA becomes double-stranded, and eventually creating a DNA that has the same sequences (U3-R-U5) at both ends (called the long terminal repeats or LTRs). The nucleocapsid protein facilitates the transfer of primers to initiate (-)cDNA and genomic strand ((+)cDNA) synthesis. Abbreviations are: R, repeated sequence; U5, unique 5' region; PBS, primer binding site; PPT, polypurine tract; U3, unique 3' region. It should be noted that all retroviruses usually have at least one PPT. HIV has two PPTs, one near the 3'end of the RNA genome and the other (the central PPT) resides near the middle of the genome [18]. For simplicity, only the 3' PPT is shown. Figure is adapted from Coffin *et al.*, 1997 [5].

position the RNA polymerase II at the correct site for initiating transcription. The viral RNA transcribed from the proviral DNA is either packaged into newly-assembly virions as genomic RNA or translated in the cytoplasm to form viral proteins [12]. The viral RNA and proteins will be transported to the plasma membrane and assembled into progeny virions. Viral maturation begins concomitantly with or immediately after budding, and the viral protease processes the Gag-pol polyprotein precursors into their mature components and eventually results in the production of infectious virus particles [12] (Fig.1.1B).

Reverse transcription serves as a critical step in retroviral replication, where the viruses use a virally encoded RT to convert its RNA genome into DNA [18]. HIV RT is a heterodimer that is composed of two subunits: p66 and p51. The p66 and p51 are 560 amino acids and 440 amno acids in length, respectively [19]. Unlike the p51 subunit which is thought to play a structural role, the p66 has a catalytic function and consists of two enzymatic activities: an RNA or DNA-dependent DNA polymerase and an RNaseH that cleaves RNA only if the RNA is part of an RNA/DNA duplex. The chemical step during the polymerization reaction requires two divalent Mg^{2+} ions [18, 19]. Near the 5' end of the HIV genome exists an 18 nucleotides long sequence named the primer binding site (PBS) and the viral reverse transcription initiates with the RNA genome being primed by a host tRNA^{lys3} at the PBS (Fig. 1.2). The RT synthesizes the (-)cDNA while the RNaseH domain of RT catalyzes the degradation of the RNA component of the DNA-RNA heteroduplex. HIV nucleocapsid chaperones the first strand transfer of the (-)cDNA to the 3' repeated (R) sequence and the polymerization of (-)cDNA continues. With the completion of (-)cDNA polymerization, (+)cDNA synthesis is primed by RNaseH-resistant polypurine tracts (PPT) of the RNA-DNA heteroduplex [5]. The final product of the reverse transcription process is a double-stranded DNA segment with each end of the DNA containing the same sequences: U3-R-U5. These are the LTRs that will serve as the substate for integrase and be the ends of the provirus after an integration event (Fig. 1.2).

1.1.2 Current HIV-1 therapies

The HIV life cycle presents many potential targets for therapeutic intervention. Currently, there are thirty nine Food and Drug Administration (FDA)-approved drugs that are available for the clinical treatment of HIV infection and they can be categorized into eight classes based on their unique molecular mechanisms: nucleoside-analog reverse transcriptase inhibitors, non-

nucleoside reverse transcriptase inhibitors, integrase inhibitors, protease inhibitors, fusion inhibitors, entry inhibitor, pharmacokinetic enhancers, and combination HIV medicines (www.aidsinfo.nih.gov). The administration of a cocktail of antiretroviral drugs have seminally facilitated the reconstitution of the immune system and reduced the morbidity and mortality associated with HIV infection in patients [20]. However, due to the high mutation rate commonly associated with RNA virus replication and frequent recombination, the generation of viral quasispecies occurs over time in HIV infected individuals. This can result in drug resistant isolates that have limited the efficacy of the current antiviral drug regimen and is worsened if there is a lack of patience adherence, which can occur due to drug toxicity and high costs. Hence, more efficient treatment strategies are required for the advancement of HIV therapy.

The discoveries of several host cellular proteins in innate immunity that are involved in inhibiting a broad range of viral pathogens (so called "restriction factors") have shed light on the development of novel anti-HIV drugs. The rationale behind this therapeutic approach is that because successful pathogens have usually developed mechanisms to counteract host defenses to establish infection, disrupting the viral counter-restriction mechanisms to fully unleash the antiviral potential of these innate host proteins would significantly decrease any possible side-effects that are associated with the traditional antiviral therapy. A host restriction factor should have at least four defining characteristics [21]. First, it must directly and dominantly cause a significant decrease in virus infectivity. Second, it must be susceptible to neutralization by viral countermeasures. Third, a restriction factor should show signatures of rapid evolution due to direct protein-protein interactions that occur through viral-host interaction. Fourth, as part of the innate immune response, the expression of a restriction factor should be strongly induced by interferons [21]. It should be noted that restriction factors are relatively rare and are different from host proteins that can negatively impact viral replication (i.e., dependency factors). Examples of the HIV restriction factors include APOBEC3s [22-25], tetherin/BST2 [26, 27], Tripartite motif-containing protein 5 alpha (TRIM5α) [28-30], SAM-domain and HD-domain containing protein 1 (SAMHD1) [31, 32], and Myxovirus resistance 2 (MX2) [33]. In this Chapter, I will focus my review on the current knowledge of the restriction factor family of APOBEC3 cytidine deaminases and their viral countermeasure, HIV Vif.

1.2 APOBEC3 family of cytidine deaminases
1.2.1 Overview

Retrotransposons and endogenous retroviruses have been genomic parasites in organisms throughout evolution and have contributed to both species evolution and disease [34]. The APOBEC (Apolipoprotein B mRNA-editing enzyme-catalytic polypeptide) family of enzymes present in their earliest form in bony fish acted as a defense against retroelements [35]. Due to expansion of retroelements through evolution there was a corresponding expansion in the APOBEC family [36, 37]. The most recent expansion in placental mammals formed the Apolipoprotein B mRNA-editing enzyme-catalytic, polypeptide-like 3 (APOBEC3) family in response to ancient pathogenic retroviruses [37, 38]. Humans contain seven APOBEC3 (A3) enzymes (A3A, A3B, A3C, A3D, A3F, A3G, and A3H) [37, 39]. The A3 enzymes act as host restriction factors to inhibit retroelement replication through either RNA binding ability or activity as ssDNA cytosine deaminases that catalyze the formation of promutagenic uracils [40-48]. Currently, A3 enzymes are primarily studied for their ability to restrict the replication of retroviruses (such as HIV) [22-25, 45, 49-53] and other viruses that replicate with a ssDNA intermediate (such has Hepatitis B Virus) [54, 55]. Restriction of the replication of these present-day viruses occurs primarily through the deoxycytidine deamination activity of A3 enzymes which results in hypermutated and inactivated viral genomes. The gene duplications that resulted in the human A3 repertoire formed two general groups of deaminases with different Zinc (Z)-coordinating domains: A3A, A3C, A3H are enzymes with a single Z-domain and A3B, A3G, A3D, and A3F enzymes with two Z-domains (Fig. 1.3A) [37]. For A3 enzymes with two Z-domains, only one domain is catalytically active, except for A3B, which may have two catalytically active domains (Fig. 1.3A) [56-59].

For HIV to successfully infect humans, it must overcome numerous physical and immunological barriers [21, 60, 61]. Within cells, HIV must overcome a network of restriction factors that are able to block specific replication steps of the virus, including A3 enzymes [21, 62]. HIV uses Vif to overcome A3 enzymes [23-25, 63, 64]. The Vif protein of simian immunodeficiency virus (SIV), the non-human primate form of the virus, has co-evolved with species-specific A3s for millions of years [65]. The HIV-1 predecessor, SIV$_{cpz}$ from chimpanzees, underwent a key evolutionary event that altered the 3' region of the *vif* gene that was essential for antagonism of human A3 function, along with further evolutionary changes in chimpanzees

8

A

A3A
A3B
A3C
A3D
A3F
A3G
A3H

Z1-type domain
Z2-type domain
Z3-type domain

Deamination
Encapsidation Domain
Vif Interaction
Processivity

B

Cytosine → Uracil

C

5'————CCC————**cDNA/(-)DNA**

↓ A3G

5'————CCU————**(-)DNA**

↓ Second strand synthesis

5'————CCU————**(-)DNA**
3'————GG**A**————**(+)DNA**

↓ DNA repair

5'————CC**T**————**provirus**
3'————GG**A**————
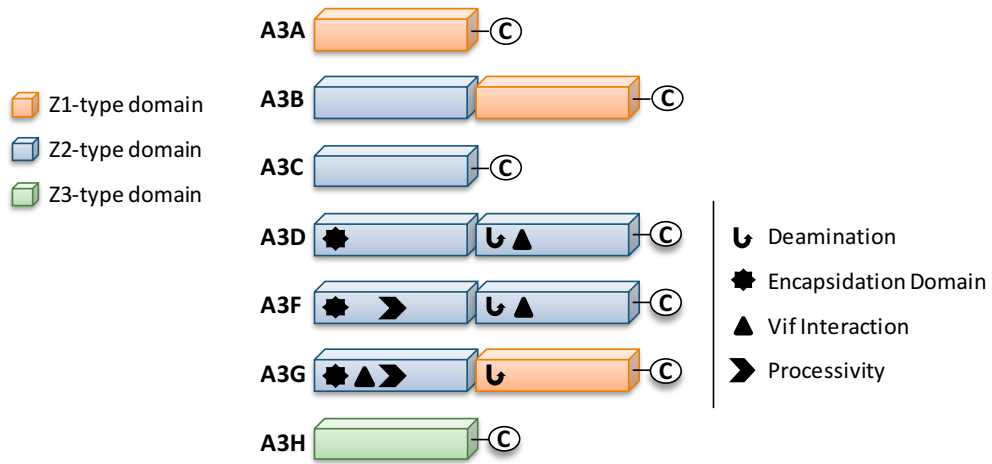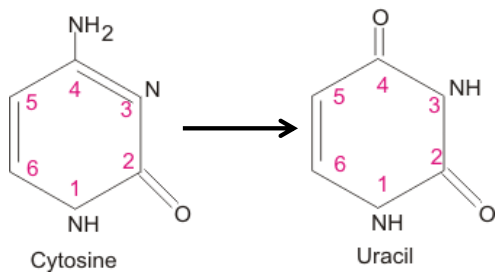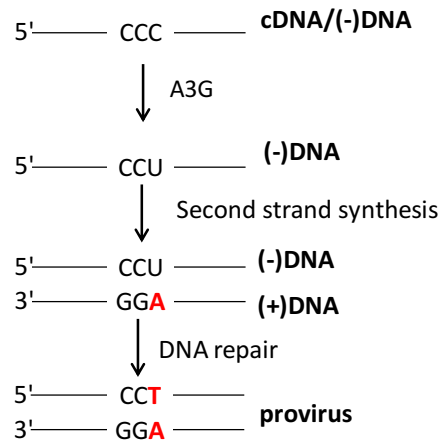
Figure 1.3 A3s catalyze the formation of promutagenic uracils in DNA through deamination reactions. (A) Zinc (Z) coordinating-type domains of human A3 enzymes. A3 enzymes coordinate Zinc through the motif H-X-E-X$_{23-28}$-P-C-X$_{2-4}$-C. The glutamate activates a water molecule to enable Zinc-hydroxide mediated nucleophilic attack to complete the deamination reaction. Deamination activity has been demonstrated for all APOBEC3 enzymes. For the enzymes with two Z-type domains that restrict HIV in CD4$^+$ T cells (A3D, A3F, and A3G), the legend depicts known biochemical functions of each Z-type domain. A common feature of A3 enzymes with two Z-type domains is the segregation of functions in the N-terminal domain (NTD) and C-terminal domain (CTD). The NTD is responsible for encapsidation and the CTD is responsible for deamination activity. Both domains can bind nucleic acids. The binding site of Vif is in the NTD for A3G and in the CTD for A3D and A3F. (B) The deamination of cytosine to uracil. DNA can undergo major structural changes as a result of thermal fluctuation. Depurination and deaminations are the two most frequent spontaneous hydrolytic chemical reactions known to create DNA damage in cells, if not repaired properly. Major types of deamination reactions include the conversion of cytosine and 5-methylcytosine to uracil and thymine, respectively, and the conversion of cytosine to uracil is shown here. Aside from spontaneous cytosine hydrolysis, uracils can also be introduced into DNA by misincorporation of dUTP instead of dTTP and through A3 cytosine deaminase catalyzed deamination of cytosine to uracil. A3 enzymes can coordinate a direct nucleophilic attack at position 4 of the pyrimidine ring. (C) A3G-catalyzed deamination reaction can lead to C/G→T/A transition mutations. A3G deaminates the cytosines on the (-)cDNA from the RNA genome of HIV. A3G prefers to deaminate the 3'C in its favored motif, 5'CCC. Upon synthesis of the (+)cDNA, instead of inserting a guanine opposite a cytosine, according to Waston-Crick base pairing, reverse transcriptase will treat the uracil as a thymine and insert an adenine and this leads to C/G→T/A transition mutation on the coding strand. Figure 1.3A is reprinted from Feng *et al*., 2014 [66].

that adapted simian immunodeficiency virus infecting chimpanzees (SIV_{cpz}) for improved infection of humans [67]. To antagonize A3 enzymes, HIV Vif must maintain the ability to physically interact with relevant A3s, host protein CBFβ for stability [68, 69], and components of the host ubiquitin ligase assembly pathway [70-72]. Ultimately Vif thwarts A3s by inducing their polyubiquitination and degradation through the proteasome (Fig. 1.4, A and B). It is thought that by disrupting the Vif-host cell interactions through novel pharmaceuticals, A3 enzymes can be used to suppress HIV. However, the natural balance of A3 enzymes and HIV must be first understood since there is evidence that HIV can take advantage of A3 enzymes to accelerate its quasispecies evolution [73, 74], although recent research suggests the inactivation of HIV by A3 is an "all-or-nothing" phenomenon and concluded that the A3-mediated hypermutation usually results in lethal mutagenesis (see below) [75, 76].

From cell culture studies, it appears that only A3A, A3D, A3F, A3G, and A3H are relevant to HIV restriction [77-81]. A polymorphism in human A3C that changes Serine to Isoleucine at amino acid position 188 (S188I) also conveyed strong antiviral activity against HIV [82]. It is not surprising that not all seven A3 members restrict HIV replication since they likely evolved to restrict different retroelement pathogens [37]. There are two paradigms of how A3 enzymes can suppress HIV. A3A present in HIV target myeloid cells can restrict replication of incoming virions through low levels of deamination and possibility another mechanism that is not yet fully elucidated [77, 79]. In CD4$^+$ T cells, A3D, A3F, A3G, A3H and A3C I188 restrict HIV $\triangle vif$ by becoming virion encapsidated in the HIV producer cell and travelling with the virion to the next susceptible cell where they catalyze promutagenic deaminations of cytosine to uracil in nascent single-stranded HIV (-)DNA (Fig. 1.4A) [78].

1.2.2 A3s that are relevant to HIV-1 restriction in CD4$^+$T cells
The A3D, A3F, A3G, A3H and A3C I188 molecules that escape Vif-mediated inhibition can restrict HIV by entering the assembling virus particle by binding RNA (HIV genome or cellular RNA such as 7SL or Y) that also interacts with the NC portion of the Gag polyprotein [83-89]. After the virus enters the next target cell A3 enzymes exert their anti-viral function during the reverse transcription process (Fig. 1.4, A and C) [45, 51, 52, 90, 91]. Although A3D, A3F, A3G, A3H and A3C I188 are localized to the cytoplasm they require encapsidation to restrict HIV and are not able to access the (-)DNA of an incoming virus (Fig. 1.4A) [78]. This may be due to the HIV capsid structure or that A3 enzymes can reside in regions of RNA
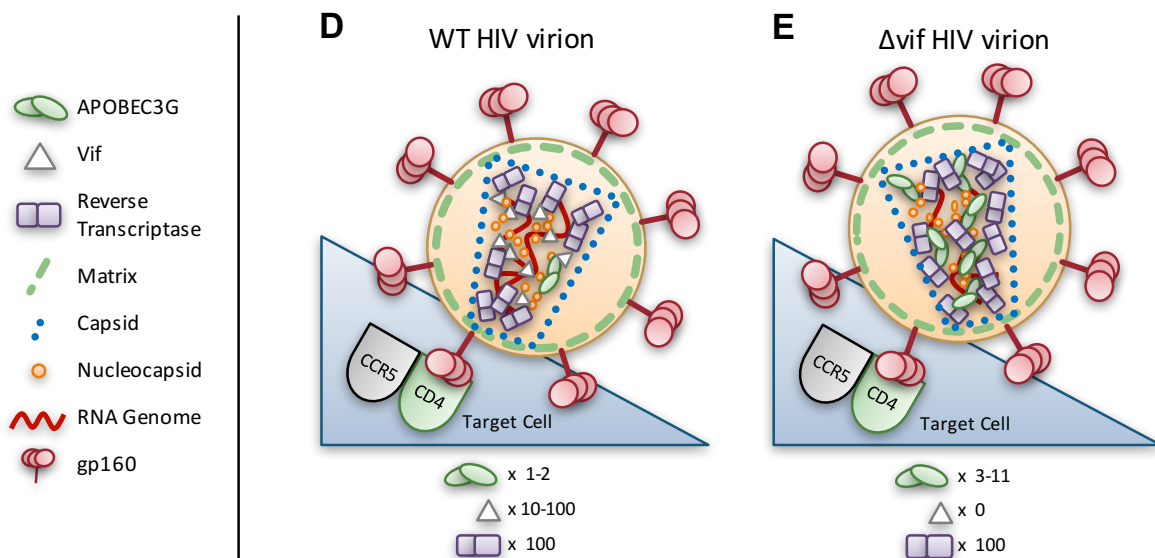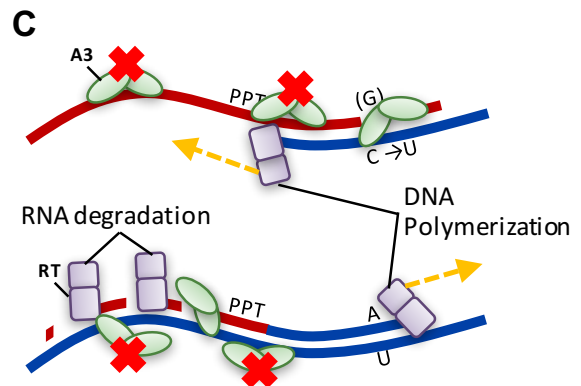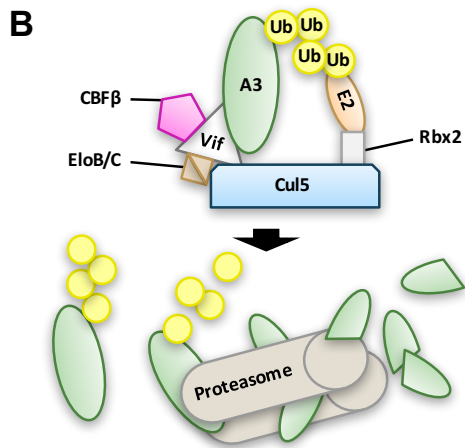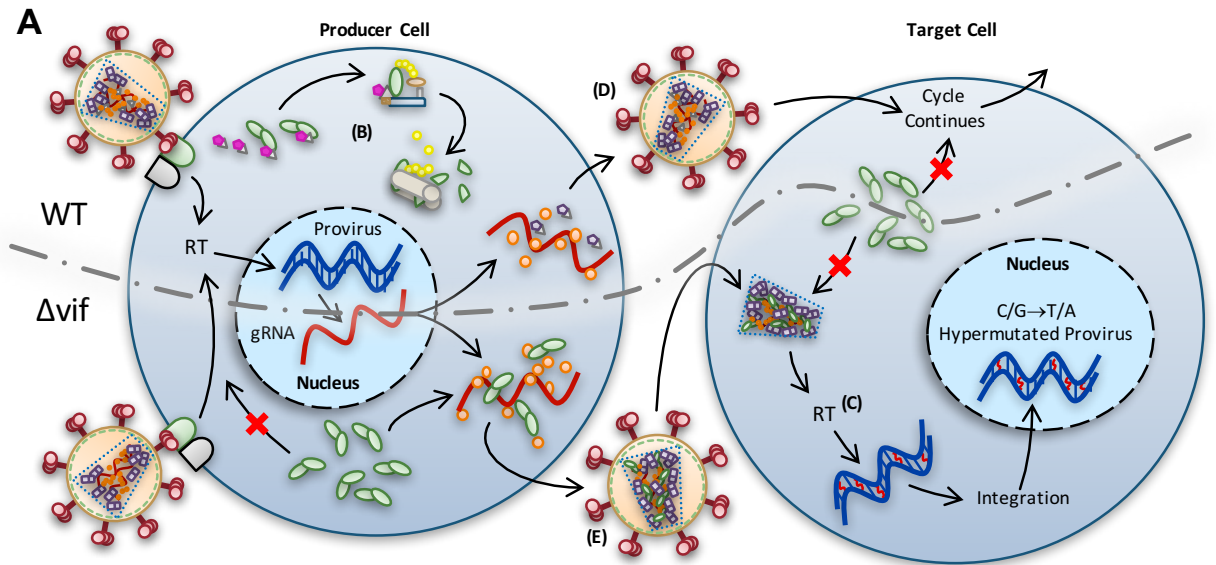
Figure 1.4 Overview of HIV restriction by A3 enzymes. (A) Sketch depicting lifecycles of wild type (WT) and ΔVif HIV (ΔVif). Each virion enters a cell that expresses A3 enzymes. In the WT virus, Vif is expressed in the cell and recruits host cell CBFβ for stability and an E3 ubiquitin ligase complex comprised of Elongin B/C (EloB/C), Cullin5 (Cul5) and Rbx2 (inset (B)). In this complex Vif acts as the substrate receptor and induces degradation of A3 enzymes. As a result, egressing virions do not encapsidate high levels of A3 enzymes and upon infection of a target cell the HIV lifecycle continues. The ΔVif HIV encapsidates A3 enzymes through an RNA and Gag interaction. In the target cell the A3 enzymes within the capsid of HIV can deaminate cytosines to uracils in nascent single stranded (-) DNA during reverse transcription (inset (C)). These uracils induce G→A transition mutations upon synthesis of (+) DNA (inset (C)). The resulting hypermutated virus can be integrated into the host genome but is functionally inactivated. A3 enzymes in the target call cannot enter the HIV capsid and are unable to restrict virus replication unless encapsidated into budding virions. (B) Detailed sketch of Vif-mediated polyubiquitination of A3G. Vif interacts with Elongin C (EloC), an obligate heterodimer with Elongin B (EloB), and Cul5. The transcription cofactor CBFβ stabilizes Vif. Cul5 binds to Rbx2, which recruits an E2 ubiquitin conjugating enzyme. Vif is the substrate receptor that recruits A3 enzymes. The ^{48}K-linked chains result in proteasomal degradation of the A3. (C) Sketch demonstrating the limited vulnerability of single-stranded (-) DNA to A3-mediated deamination that is imposed from the dynamics of reverse transcription. Reverse transcriptase is abbreviated as RT. HIV contains two polypurine tracts (PPT) that are used as a primer for (+)DNA synthesis. In the figure, only one PPT is depicted. (D-E) Sketches depicting the stoichiometry of major virion components for a (D) WT and (E) ΔVif HIV virion. Figures correspond to insets (D) and (E) in panel (A). (D) Low amounts of A3 may escape Vif mediated degradation and become virion encapsidated (approximately 1-2 molecules of A3G/virion). (E) A ΔVif HIV cannot induce degradation of A3 enzymes which results in the encapsidation of A3 enzymes through an interaction with RNA and Gag. Approximately 3-11 molecules of A3G can be encapsidated from the virus producer cell. (D-E) Stoichiometry values for virions were obtained from [5, 92-95]. Figure reprinted from Feng *et al*., 2014 [66].

processing, e.g., stress granules or P-bodies, where they may have a role in sequestering human retrotransposon RNA to prevent transposition [42, 96-98]. A3 enzymes are Zinc-dependent deaminases and they catalyze the conversion of cytosine to uracil in ssDNA substrates. The reaction requires the activation of water by a Zn^{2+} coordinated by the enzyme. A glutamic acid in the enzyme active site primes the nucleophilic attack on the C4 position of the pyrimidine ring, followed by the removal and protonation of an amino group (NH_2) and this leads to the release of ammonia (NH_3) and formation of uracils as products (for review, see [99]) (Fig. 1.3B). Each ssDNA deaminase acts within a preferred di- or tri- nucleotide substrate motif. For A3G this is 5'CCC or 5'CC (3'-end C preferred for deamination) [91] (Fig. 1.3C). A3D, A3F and A3H and A3C I188 deaminate 5'TTC or 5'TC and A3D can also deaminate 5'GC motifs in proviral DNA [49, 52, 100]. Since A3 enzymes are ssDNA deaminases, deamination activity is restricted to the (-)DNA strand (Fig. 1.3C) [91]. The cytosine (C)→ uracil (U) deaminations catalyzed on the (-)DNA strand become guanine (G)→adenine (A) mutations when RT uses uracil as a template during (+)DNA strand synthesis (Fig. 1.3C) [91]. The resulting 'hypermutation' of the provirus results in inactivation of HIV (Fig. 1.4A) [22-24]. Although it is known that many proviral genomes undergo successful integration with these hypermutations [101], some preintegration complexes containing uracils may be degraded by host DNA repair mechanisms, although there is no consensus regarding the extent to which this occurs in cells [102, 103]. In cell culture, it has been found that the aberrant HIV proteins that may be produced from these proviral genomes can act as a source of HIV antigens due to their misfolding and processing through the proteasome, which can facilitate immune recognition of HIV [104]. The aberrant HIV protein products produced due to A3-induced missense and nonsense mutations could also facilitate the cytotoxic T lymphocytes (CTLs) or natural killer cell-mediated clearance of infected cells [104, 105], although other reports found that A3-mediated mutations can reduce CTL recognition and induce the generation of CTL response escape mutants of HIV [106-109].

Although the majority of A3 actions are repressed by Vif in HIV infected individuals (Fig. 1.4, A,B,D), clinical studies have found that individuals with an inherent ability to express a high level of A3G mRNA are less likely to become infected with HIV or progress from HIV to AIDS and that the presence of hypermutated proviral genomes correlates with high $CD4^+$ T cell counts [110-114]. Other A3 enzymes have not been extensively examined in this regard [100, 115].

However, there is evidence of deaminations in HIV genomes recovered from infected individuals due to C/G→T/A mutations in a sequence context that indicates deaminations by A3 enzymes other than A3G do occur [100, 111, 113, 114]. For example, a study found that in HIV infected individuals suggests there is approximately an equal split between mutations occurring in the 5'TC and 5'CC contexts on (-) DNA [100]. However, there is difficulty parsing out the effect of A3F, A3D, and A3H based on their mutation patterns since they all recognize the minimal dinucleotide 5'TC and are more promiscuous than A3G in regards to trinucleotide target site preference [49, 52, 78]. Nonetheless, it has become clear that despite some evidence that A3G has more mutagenic potential than other A3 deaminases, it is not acting alone against HIV [81, 100, 116] and proviruses with GG and GA mutation signatures most likely result from copacking of multiple A3s and comutation of the same viral genomes in a single round of replication [117] (see below). The HIV genomes mutated through A3 catalytic activity are also subject to the pressure of purifying selection. This selection pressure results in mutated and inactivated genomes being highest in integrated proviral DNA and lowest in circulating viral RNA [101], which means only a small proportion of hypermutated genomic RNA can be packaged into virions. Further, integrated proviruses that are inactivated by stop codons in the *gag* gene may still be rescued by dual infection of a cell by HIV quasispecies and complementation of Gag function [101]. Recombination within virions by RT template switching can result in "reactivation" of inactivated viral genomes [101, 118]. As a result, A3-mediated mutagenesis is effective, but the complete inactivation of HIV in an infected individual is potentially a long term process that is likely to require multiple rounds of exposure to viruses [119].

HIV populations generally exhibit high levels of genetic variation and this allows the virus to easily acquire drug resistance or escape immune surveillance [120]. The viral genome diversity is generated due to the fact that HIV RT lacks proofreading activity [121] and that HIV copackages two viral RNA genomes per virion and recombination during viral replication occurs frequently [122, 123]. The deamination activity of A3 enzymes can also contributes to HIV genetic variation through two mechanisms, namely by introducing sublethal levels of G-to-A mutations in HIV (i.e., sublethal mutagenesis) and through copackaging and recombination, lethally mutated genome can be rescued by superinfecting the hypermutated provirus-containing cells with a replication-competent wild-type virus [76]. The effects of A3-induced G-to-A mutations on HIV mutation rate, recombination and the emergence of drug resistance are still

under debate. Despite several previous studies that reported A3 enzymes contribute to viral genetic variation and evolution through sublethal mutagenesis [73, 74, 118, 124, 125], other studies proposed that A3-mediated hypermutations usually lead to lethal mutagenesis [75, 76, 126-128]. Recently, Delviks-Frankenberry *et al.* [76] found that the contribution of A3G or A3F hypermutation made to HIV sequence variation and evolution is minimal and is significantly small when compared to mutations induced by HIV RT. Also, A3G or A3F-mediated mutations do not appear to affect the rate of recombination. Their study [76] reconciles that A3-induced sublethal mutagenesis being reported by different groups [73, 74, 118, 124, 125] likely exists, but occurs at a low level, therefore its contribution to genetic variation will likely get titrated out with each successive replication cycle.

1.3 A3-mediated restriction of HIV-1
1.3.1 A3G-mediated restriction of HIV-1
1.3.1.1 Deamination-dependent restriction of HIV-1 by A3G

Since A3G was the first A3 enzyme discovered it has been more widely studied and there is a wealth of information available to explain how A3G enacts its role as a restriction factor (over 900 publications in PubMed). There are two key steps that A3G must complete to be an efficient restriction factor. First, the enzyme must be available for binding RNA that will become virion encapsidated through interaction with the NC portion of Gag [83-89, 129-131]. Second, it must have a mechanism to scan the nascent HIV (-)DNA that is available for a finite period of time for potential cytosines that it can deaminate (Fig. 1.4C) [93, 132, 133].

A3G exists in cells as a high molecular mass that is bound to RNA and other proteins in stress granules and P-bodies [42, 96, 98, 134]. However, only newly synthesized A3G that has not associated with host RNAs in these cytoplasmic structures appears to bind the RNA that is also bound by HIV Gag and therefore encapsidated into virions [135]. A3G requires oligomerization to bind these RNAs effectively in cells and become virion encapsidated [136-138]. *In vitro*, oligomerization mutants of A3G can bind many RNAs with less than a 3-fold difference from wild type A3G [139, 140]. The RNA binding and oligomerization of A3G is primarily mediated by the N-terminal domain (NTD) and the NTD is solely responsible for virion encapsidation of A3G (Fig. 1.3A) [56, 57, 138, 140]. The NTD residues [124]YYFW[127] on predicted loop 7 mediate the dimerization of A3G (Fig. 1.3A) [138, 140]. A3G is primarily a dimer in solution and when A3G binds RNA or DNA it oligomerizes into tetramers and

A A3G NTD model
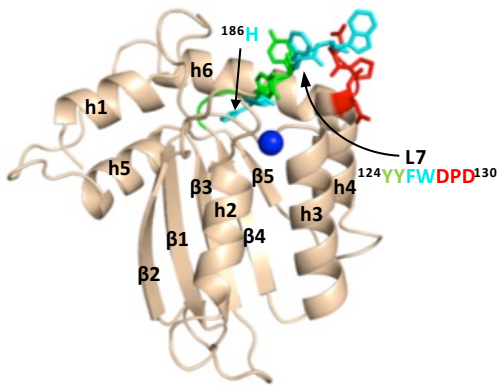
B A3G CTD
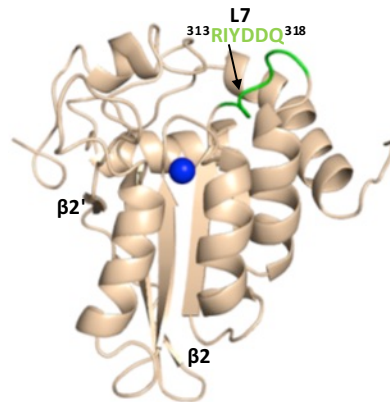
C A3F NTD model
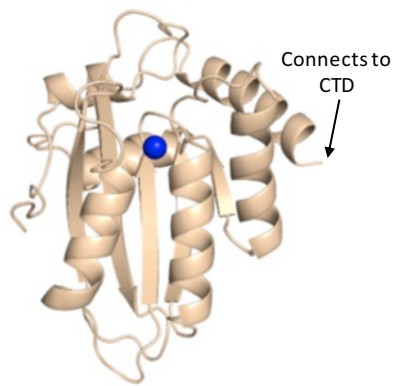
D A3F CTD
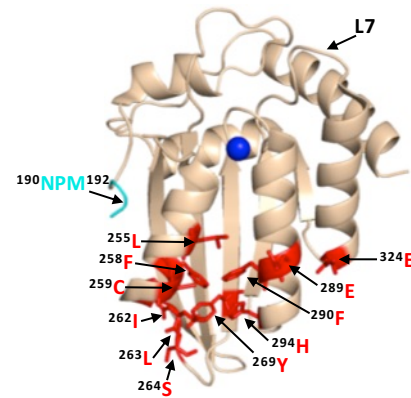
E A3D CTD model
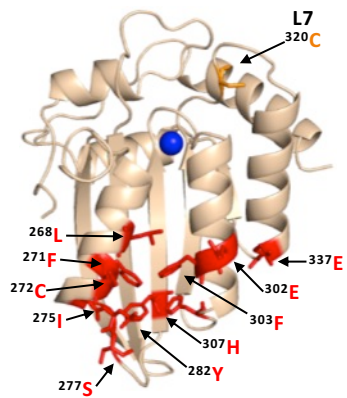
F A3H Hap II model
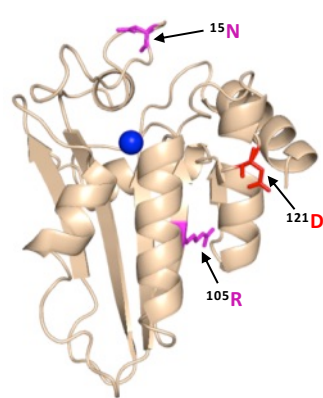
Figure 1.5 Structures of A3 enzymes. A3 enzymes have a basic structure in each Z-type domain that is comprised of a five stranded β-sheet core surrounded by six α-helices. Numerical assignments to β-strands and α-helices are superimposed in panel (A). Zinc atoms are shown as blue spheres. (A) Model of the N-terminal domain (NTD) of A3G. Loop 7 (L7) of the A3G NTD is a central structure in its anti-HIV function. Highlighted on L7 are the residues important for interaction with Vif (red, [128]DPD[130]), oligomerization/virion encapsidation (green and cyan, [124]YYFW[127]), and jumping component of A3G processivity (cyan, [126]FW[127]). Helix 6 (h6) is adjacent to L7 and contributes to the sliding component of A3G processivity, particularly [186]H (cyan). The predicted model of the A3G NTD was obtained by using the automated SWISS-MODEL program using the homologous A3G C-terminal domain (CTD, PDB: 3IQS). (B) The A3G CTD (PDB: 2KEM) is the catalytic domain of A3G and has a discontinuous β2 strand forming a loop-like buldge between the β2 and β2' strands. A3G tetramerization is mediated by L7 residues [313]RIYDDQ[318] (green). The L7 residues in all A3 enzymes also determines the preferred deamination motif. (C) Model of the NTD of A3F was obtained by using the automated SWISS-MODEL program using the homologous A3C structure (PDB: 3VM8). The end of h6 connects the NTD to the CTD and contains an [190]NPM[192] motif. This NPM motif is found only in A3D and A3F. (D) The A3F CTD (PDB: 4IOU) is the catalytic domain of A3F and interacts with Vif. Residues that interact with Vif across Helix 2, 3, 4 and β-strand 4 are shown in red. Also shown on this structure is the deamination motif specificity loop (L7) and the [190]NPM[192] motif. The structure illustrates the kinked orientation introduced by the Pro in the [190]NPM[192] motif, which blocks the sliding function of A3F. (E) Model of the CTD of A3D was obtained by using the SWISS-MODEL program using the homologous A3F structure (PDB: 4IOU). Residues that interact with Vif across Helix 2, 3, 4 and β-strand 4 are shown in red. The [320]C residue on L7 that influences A3D activity is shown in orange. (F) Model of A3H Hap II showing residues that interact with Vif and cause haplotype instability. In A3H Hap II, [121]D (red) on predicted h4 mediates an interaction with Vif. In A3H Hap I the R105G mutation induces protein instability (magenta). In A3H Hap III and IV, the deletion of ^{15}N induces protein instability (magenta). Figures were made using PyMOL (The PyMOL Molecular Graphics System, Version 1.5.05, Shrödinger, LLC.). Figure reprinted from Feng *et al*., 2014 [66].
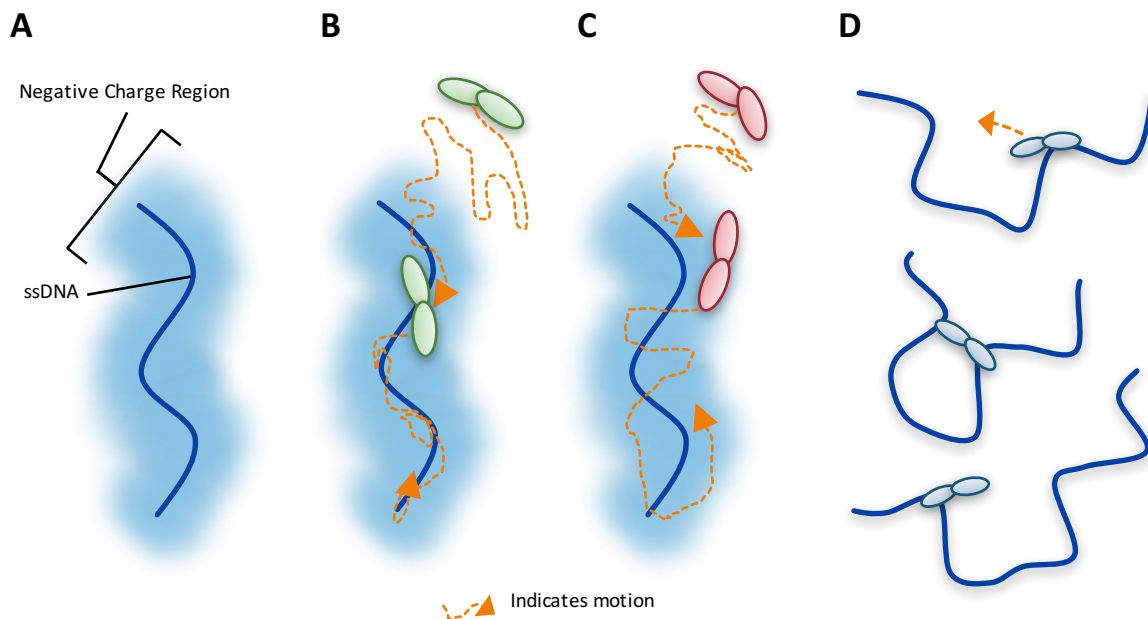
Figure 1.6 Illustration of DNA scanning by facilitated diffusion. (A) Sketch of DNA showing the negative charge region of DNA important for facilitated diffusion of A3 enzymes. (B) Sketch depicting a 1-dimensional DNA scanning path by sliding. Dotted line indicates path of enzyme. Sliding enables an in-depth search of local areas of a substrate. (C) Sketch depicting a 3-dimensional DNA scanning path by jumping. Dotted line indicates path of enzyme. Jumping enables larger translocations on DNA substrates, but lacks a local search process. The microdissociations of the enzyme from the DNA that occur when the enzyme jumps does not leave the negatively charged domain of the DNA so the enzyme has a higher likelihood of reassociating with the same DNA substrate than diffusion into the bulk solution. (D) Sketch depicting a 3-dimensional DNA scanning path by intersegmental transfer. Intersegmental transfer enables larger translocations on DNA substrates, but lacks a local search process. An enzyme with two DNA binding domains binds two regions of DNA simultaneously before dissociating from one region to move to another. Figure reprinted from Feng *et al*., 2014 [66].

higher-order structures through C-terminal domain (CTD) residues [313]RIYDDQ[318] on loop 7 (Fig. 1.5B) [140-142]. It is essential for A3G to enter the inner capsid of the virion to restrict HIV. Within the capsid, A3G can associate with the ribonucleoprotein complex and access the (-)DNA as it is synthesized. A3 enzymes that cannot encapsidate within the HIV capsid, e.g., A3A and A3C, are unable to restrict HIV replication in CD4[+] T cells [78, 143]. However, accessing the ribonucleoprotein complex does not guarantee the ability to restrict HIV. Since the (-) DNA is only available for a finite period of time due to HIV containing two polypurine tracts (PPT) that are used to prime (+) DNA synthesis, A3 enzymes require an efficient mechanism to search for cytosines (Fig. 1.2 and Fig. 1.4C). Complicating the search is that the (-) DNA contains pieces of annealed RNA due to the endonuclease activity of the RT associated RNaseH (Fig. 1.2). A3G binds RNA/DNA hybrids less well than ssDNA and encountering these obstacles on the substrate can induce A3G to dissociate from DNA [141, 144].

Unraveling the mechanism by which A3G locates its target and catalyzes its deamination is of pivotal importance for understanding the mechanistic basis of proviral hypermutation. The mechanism with which an A3 enzyme scans non-target DNA in search of its deamination motif is a determinant in its catalytic efficiency [133, 139]. DNA scanning is described by the term processivity and is defined as the ability of an enzyme to catalyze multiple events in a single enzyme-DNA substrate encounter. Enzymes that do not use an energy source for movement on DNA use a mechanism termed facilitated diffusion to efficiently search DNA [145, 146]. This is a mechanism where the enzymes, subject to Brownian motion, move randomly on DNA. Since DNA binding proteins are usually positively charged, the negative charge of the DNA facilitates the enzyme movement through electrostatic interactions [145, 146]. A3G is a positively charged enzyme (charge of +6.5 at pH 7) and processively scans ssDNA by facilitated diffusion (Fig. 1.6) [93, 132]. This mode is distinct from an enzyme that acts on DNA distributively, where only one catalytic event occurs before the enzyme disengages from the substrate [147]. Facilitated diffusion can involve a variety of movements such as 1-dimensional (D) sliding (Fig. 1.6B) or 3-D movements such as hopping/jumping (Fig. 1.6C) or intersegmental transfer (Fig. 1.6D) [148]. Hopping and jumping describe small micro-dissociations and re-associations with the same DNA strand without diffusion into the bulk solution (Fig. 1.6C) [146]. Intersegemental transfer involves a two-step mechanism where an enzyme with two DNA binding sites binds a second site before releasing the first site (Fig. 1.6D) [146]. Facilitated diffusion works best when both 1-

20

and 3- D movements are used to enable local scanning of a small segment of DNA by sliding (<20 nt) and movement to distal regions to restart the local scanning process [139, 148]. These distal movements do not cause the enzyme to leave the DNA and enter into the bulk solution because the charged surface of the DNA keeps the enzyme within the domain of the DNA (Fig. 1.6C) [146]. Using different methods A3G has been found to scan ssDNA by 1-D sliding motions and 3-D jumping motion [132, 149, 150]. However, one study has found that A3G moves by 3-D intersegmental transfers [93]. The efficiency imparted by a 3-D movement in the specific case of A3G during reverse transcription is that it provides a means of overcoming the DNA/RNA hybrid barrier [93, 139]. Clusters of A3G-induced deaminations indicative of processive sliding movements have been found in integrated proviral genomes [151] and A3G mutants unable to undergo a local searching process by 1-D sliding, such as H186R and A3G with an [195]NPM[197] insertion have decreased mutagenesis during *in vitro* reverse transcription or in HIV proviral genomes [133, 139]. Further, an A3G mutant F126A/W127A that cannot jump has decreased mutagenesis during *in vitro* reverse transcription [139]. These data demonstrate that neither movement alone can enable high levels of A3G-induced mutagenesis. Interestingly, the F126A/W127A mutant is monomeric, suggesting that the oligomeric state of A3G plays a role in efficient restriction of HIV not only by facilitating virion encapsidation but also by facilitating the DNA scanning process [138-140]. The processivity determinants of A3G reside on predicted loop 7 and helix 6 of the noncatalytic NTD (Fig. 1.3A) [139]. Thus, despite a lack of catalytic activity, the NTD contributes to A3G deamination activity by mediating the processive scanning mechanism.

1.3.1.2 Deamination-independent restriction of HIV-1 by A3G

A3G primarily restricts HIV replication through its deamination activity. However, there have been numerous reports of an ability of A3G to physically inhibit other processes of HIV such as RT polymerization [152-157], NC strand annealing [158, 159], and proviral DNA integration [156, 160]. These processes do not occur in isolation from deamination, nor do they restrict HIV better than deamination alone [156, 157]. I will focus the discussion on inhibition of RT polymerization since this is the most prevalent mechanism studied.

Early studies of A3G-mediated restriction of HIV proposed that transiently overexpressed wild type A3G and deamination null mutants of A3G could inhibit HIV proviral DNA formation [23, 161]. The initial mechanism proposed was that A3G binds the HIV genomic RNA which

impedes RT [152]. This has been confirmed in multiple reports using cellular and biochemical experiments [155, 156, 162]. However, the physiological significance of these processes is difficult to reconcile since results from cell-based experiments cannot be replicated when A3G is stably expressed, suggesting that overexpression of A3G induced experimental artifacts due to excessive packaging of A3G in virions [151, 163-165]. Importantly, studies that used deamination null mutants of A3G to show that deamination ability is required for restriction of HIV [163-165] should be considered in conjunction with data showing that the A3G E259Q catalytically inactive mutant is not a true proxy for A3G [153, 155]. A3G E259Q binds RNA less well than A3G and this results in less inhibition of RT *in vitro* and in cells [153, 155].

Nonetheless, it is clear that the ability of A3G to inhibit RT is highly dependent on A3G concentration and the primer/template [151, 155, 163-165]. The initiation of DNA synthesis from an RNA primer on an RNA template is the least efficient type of polymerization activity of RT [166]. Accordingly, *in vitro,* low levels of A3G can best inhibit RT-mediated primer initiation at this step by competing for substrate with RT [152, 155]. In contrast, on a DNA primer and DNA template, A3G could at most inhibit initiation of synthesis by RT 2-fold under single turnover conditions and could not block RT from binding the primer/template, but merely lengthened the time RT required to find the free 3'OH [155]. These data are in agreement with a computational study that suggests an A3G-mediated deamination-independent mode of HIV restriction contributes <1% of the restriction capability of A3G [167]. Although studies have shown that a peptide of A3G can interact with RT and inhibit RT-mediated DNA synthesis, it is unlikely that this mechanism is a physiological way to inhibit reverse transcription since in △Vif virions, only approximately 7 A3G molecules are encapsidated whereas there are approximately 100 RT molecules (Fig. 1.4E) [5, 92, 154]. In the presence of Vif, there is only an estimated 1-2 molecules of A3G per virion (Fig. 1.4D) [93], emphasizing the importance of deamination over a deamination-independent mechanism. A single molecule of A3G could inactivate an HIV provirus through cytosine deamination whereas the deamination-independent mechanism is much more concentration dependent [151, 155]. Single-molecule studies have brought forth the model that A3G oligomers can act as a road-block for HIV [162]. An additional component to this "road-block" model is that A3G can change the structure of DNA and RNA, presumably by wrapping or bending the substrate which may enable it to compete with RT for substrate binding and delay location of primer termini by RT [141]. Notably, existence of a deamination-

independent mode of HIV inhibition has been observed *ex vivo* in primary cells [157], but requires further research as to the significance of this mode of inhibition during an HIV infection.

1.3.2 A3F-mediated restriction of HIV-1
1.3.2.1 Deamination-dependent restriction of HIV-1 by A3F

Approximately 2 years after the discovery of A3G, A3F was discovered to also exhibit restriction activity against HIV [49-51]. Sequenced HIV proviral genomes were known to contain G/C→A/T transition mutations in 5'CC and 5'TC contexts in the (-)DNA [168-170], and A3F was found to contribute to the transition mutations in the 5'TC context. Of these initial studies demonstrating A3F activity active against HIV [49-51, 171], all except one [172] found that A3F restriction activity was equivalent to A3G restriction activity. It has since been shown that overexpression of A3 enzymes can result in excessive packaging into HIV virions and result in artifacts of HIV restriction [92, 164, 165]. Yet even after 10 years of studying A3F, multiple groups still find different activities of A3F against HIV that cannot be attributable to overexpression, but perhaps different experimental systems and techniques [78, 81, 133, 173, 174]. However, A3F must exert a restriction pressure on HIV since Vif maintains an interaction with A3F that is distinct from A3G in order to induce A3F degradation [175]. As with A3G, for A3F to be an effective restriction factor against HIV, it must be encapsidated with the ribonucleoprotein complex within the capsid and effectively search for cytosines on the heterogeneous (-)DNA substrate [133].

A3F encapsidates into HIV virions through an interaction with RNA, but packages more efficiently than A3G into the core of HIV particles [131, 172, 176]. By resolving HIV capsids on a sucrose gradient to observe whether A3F and A3G partition with the RNA and enzymes or the p24 capsid protein, Song *et al*. [176] found that more A3F specifically associated with the ribonucleoprotein complex, in comparison to A3G. A3F binds RNA with 7-fold higher affinity than A3G [133], which may enable it to package more specifically within the capsid [176]. Furthermore, A3F has been shown to bind double-stranded DNA with a higher affinity than A3G and maintain an association with the pre-integration complex of HIV as it enters the nucleus through its high affinity nucleic acid binding [177, 178]. The original functions of A3s were against unknown retroelements and A3F and A3G may have specifically evolved to inhibit different ones where these different nucleic acid binding abilities were beneficial. Although some reports show A3F can be less effective than A3G in restricting HIV [81, 133, 173, 174,

176], it cannot be concluded that it does not suppress HIV or impose selective pressure on HIV. It is not only because of data showing A3F can effectively restrict HIV in spreading infections [78, 80], but also because an HIV lab strain with tandem stop codons in Vif (from HIV NL4-3) will revert back to expressing a functional Vif in the presence of A3F [179]. This does not occur when A3G is used in the same type of forced evolution experiments [180, 181]. The HIV evolves to overcome A3G restriction, but does so by acquiring a 5'UTR mutation to make HIV RNA transcription more efficient and altering the cell cycle through a Vpr mutation [180, 181]. Together these mutations result in more virus particles being produced. Presumably since A3G has less specific packaging in the capsid than A3F, this strategy titrates out the ribonucleoprotein-packaged A3G, enabling the HIV to escape high levels of mutagenesis. These data illustrate that A3G and A3F exert a distinct selective pressure on HIV due to distinct biochemical properties and that A3 packaging into virions is a necessary but insufficient step to ensure efficient HIV restriction.

To further understand why A3F-mediated restriction of HIV may be different than for A3G, Ara *et al*. [133] undertook a biochemical study of A3F in comparison to A3G to identify biochemical differences between these enzymes that could account for differences in restriction efficiency. They found that in contrast to A3G, A3F used only 3-D jumping motions to scan ssDNA. This made the DNA scanning mechanism inefficient since A3F could translocate between many ssDNA regions and overcome intervening RNA/DNA hybrid regions, but lacked a local search mechanism to examine ssDNA regions for its 5'TC motif (Fig. 1.6). The A3F sliding movement is blocked by a ^{190}NPM^{192} motif in the connection domain between the NTD and CTD since mutagenesis of this motif to ^{190}NGM^{192} enabled A3F to slide [133]. Bohn *et al*. [182] reported an A3F CTD structure that includes the ^{190}NPM^{192} sequence and showed that it is a kinked region of the loop structure. Since an A3F ^{90}NGM^{192} mutant was able to slide, the data suggest that the rigid ^{191}P residue primarily blocks sliding [133]. However, imparting sliding movement to A3F through the ^{90}NGM^{192} mutant did not increase A3F HIV restriction efficiency because the jumping movements of A3F differed from A3G and were dominant over sliding which maintained an inefficient search of ssDNA [133]. The differences in DNA scanning between A3F and A3G were shown to be relevant to HIV restriction since A3F was 4-fold less effective in restriction of HIV than A3G in a single-cycle infectivity assay [133]. Of note, A3F was also shown to have a 100-fold lower specific activity than A3G [133] but this was not

thought to contribute to differences in restriction efficiency since studies with different A3G and A3F mutants showed that mutagenesis efficiency correlated with the efficiency of the ssDNA scanning mechanism, not the specific activity. This is likely because RT polymerization and RNaseH activity limit the (-)DNA substrate available [183]. The study by Ara *et al*. [133] is in agreement with studies where A3F has not been as effective as A3G in restriction of HIV [81, 173, 174], although A3F was found to be as restrictive to HIV replication as A3G in other reports [78, 80]. Despite A3F being considered in some reports to be less efficient than A3G as an HIV restriction factor when considered side by side, this is far from the natural mechanism of these enzymes in which they act in concert [80, 100].

It is of note that Zennou and Bieniasz reported that per mutation, A3F was less likely to inactivate HIV than A3G [172]. This was later found to be because the 5'CC motif of A3G overlaps with the only Trp codon (5'TGG/ACC) and results in a stop codon upon deamination of either cytosine in the motif. In contrast, A3F induced mutations largely result in missense mutations which may or may not inactivate the encoded protein. The A3G 5'CC motif also overlaps with Gly codons and in the HIV *prot* mutations at these Gly results in more nonconservative mutations and gene inactivation than A3F-induced missense mutations that primarily cause the conservative mutation of Glu to Gln [133]. The determinant for motif specificity is loop 7 in the CTD (Fig. 1.5, B and D) [184-187]. This loop can be grafted into different A3 enzymes to change site specificity [185]. However, the consequences of deamination-mediated restriction can be independent from inducing amino acid changes. A3F and A3G may be able to block proviral integration through deoxycytidine deaminations that result in aberrant processing of the proviral DNA ends by HIV integrase and inhibition of plus-strand DNA transfer by reducing the efficiency of primer tRNA removal [163, 177].

1.3.2.2. Deamination-independent restriction of HIV-1 by A3F

For many years, A3F was thought to have a stronger deamination-independent mode of inhibiting HIV than A3G [188, 189]. Unlike A3G, the mechanism of deamination independent "activity" was not extensively studied, but was presumed to be due to inhibition of RT polymerization. A computational study has found that A3G and A3F rely differentially on their deamination independent modes of HIV restriction with A3G only having the deamination independent mode contributing to <1% of its restriction activity whereas for A3F this value was approximately 30% [167]. However, two studies using stably expressed A3F and A3F catalytic

mutants C280S/C283A and E251Q demonstrated no inhibition of RT, suggesting that previous results were influenced by A3F overexpression artifacts [173, 190]. Another study showed that A3F can inhibit HIV integration by reducing 3' processing of viral DNA at the U5 and U3 ends by integrase [177]. Using a catalytic mutant of A3F (E251Q), the study found that inhibition of integration was decreased approximately 2-fold from that of wild type A3F suggesting that catalytic activity is in part required to produce the aberrant U5 and U3 ends [177]. Thus, there appears to be consensus that despite the potentially inefficient mutagenic activity of A3F in some studies, the deamination activity of A3F is still dominant over the deamination-independent activity. Further, if a deamination-independent mode of HIV inhibition functions in cells, it may be the inhibition of integration rather than RT [177].

1.3.3 A3D-mediated restriction of HIV-1

A3D was first characterized in 2006 to restrict HIV replication in single-cycle infectivity assays and to be suppressed by Vif, suggesting that it posed a restriction pressure on HIV[52]. Further evidence of this was that HIV proviral genomes showed evidence of deaminations in both the contexts of 5'CC, 5'TC, and 5'GC [52]. A3D was found to deaminate in the 5'TC and 5'GC contexts which were unique from A3G and A3F that maintain less promiscuous deamination motif preferences [52]. A3D also forms multimers through an RNA intermediate in cells with a similar profile as A3G [191]. In a clinical study of HIV infected individuals, A3D was found to be upregulated in both Elite Controllers (low viral load) and in Non-Controllers (high viral load) but was down-regulated in response to successful anti-retroviral treatment, indicating that A3D is indeed part of an immune response to HIV [192]. However, the restrictive activity of A3D appears less than A3G and A3F in single cycle infectivity assays in cell lines [52, 193] and spreading infections in primary human cells [81]. Further, A3D represents the most divergent A3 enzyme in the lineage of chimpanzee to human and the activity of A3D has decreased from chimpanzees to humans [47]. Chimpanzee A3D induces more hypermutation of HIV than human A3D, despite equivalent packaging into virions [47]. This was attributed to differences in loop 7 of the CTD [47, 193]. One report found a single amino acid in the CTD loop 7, C320, that suppressed A3D antiviral activity (Fig. 1.5D) [193]. If the C320 was replaced with a Tyr, as in A3F, the activity of A3D could be increased by more than 20-fold [193]. In contrast, endogenous A3D from the T cell line CEM2n appears to have activity against HIV-1 that is similar to A3F [80]. Using a series of *A3 null* backgrounds or A3 knockdowns, Refsland

et al. [80] found that the HIV-1 proviral hypermutation pattern at 5'CC and 5'TC sites was induced at comparable levels by the combined action of A3G and A3F or A3G and A3D, suggesting a redundancy in the HIV-1 restriction mechanism.

1.3.4 A3H-mediated restriction of HIV-1

A3H was originally identified as not being able to restrict HIV replication due to low steady-state protein levels in cells, despite normal mRNA expression [52, 194]. However, when A3H was recombinantly expressed in *Escherichia coli* (*E. coli*) it could mutate the *E.coli* genomic DNA. In later studies, it was realized that A3H exists as multiple haplotypes in the human population (Hap I-VII) with different stabilities in cells and HIV restriction capabilities (Table 1) and the original A3H tested was an unstable form (Hap I) [45, 195]. The unstable Hap I is the most prevalent form of A3H in the population (Table 1) [196], but is able to restrict HIV infection by approximately 2-fold when transiently overexpressed in cell lines [45, 195, 197, 198]. Two amino acid polymorphisms, ^{105}G and Δ^{15}N, can independently contribute to the instability of A3H (Table 1). A3H Hap I is unstable due to a Gly at position 105 [45]. An A3H Hap I G105R mutant (later identified as Haplotype VII, Table 1) renders the A3H stable in cells and imparts strong anti-HIV activity [45, 195]. Other unstable A3H haplotypes (III and IV) have the Δ^{15}N in combination with another polymorphism (Table 1) [45, 195]. A3H polymorphisms have been shown to be associated with HIV susceptibility in different ethnic groups [199, 200], as the haplotype frequency of unstable A3H Hap III is low and that of the stable A3H Hap II is high in HIV long term non-progressors, indicating the expression of a stable A3H haplotype may help in delaying disease progression to AIDS.

It is not known biochemically why these A3H haplotypes are unstable, but comparative modeling of A3H with the structure of a related family member APOBEC2 shows that amino acid 105 is in a β-strand within the enzyme core five stranded β-sheet, suggesting that a R105G mutation could destabilize the core structure (Fig. 1.5F). The Δ^{15}N is predicted to be within a loop structure (Fig. 1.5F) so it is difficult to predict the reason for the instability in this undefined region, but it is known from studies with A3F that deletions to a loop that connect the NTD and CTD cause protein instability [133], suggesting that the A3H loop may need to be of a specific length for proper protein folding. Although different haplotypes (II, V and VII) have been reported to exist in the population as stable forms that are able to restrict HIV (Table 1), in this review we focus only on A3H Hap II (A3H Hap II), which has been the most highly studied.

Table 1. Summary of A3H Haplotype features

A3H Haplotypes	Polymorphic amino acid residues					Antiviral activity in cell culture	Protein stability	Haplotype frequency	References
	Δ15	18	105	121	178				
Hap I	N	R	G	K	E	Partial	No	0.526[a] 0.308[b]	[45, 78, 100, 194, 195, 197, 198, 201]
Hap II	N	R	R	D	D	Yes	Yes	0.061[a] 0.265[b]	[45, 78, 100, 197, 198, 201]
Hap III	Δ	R	R	D	D	No	No	0.070[a] 0.114[b]	[45, 100, 198]
Hap IV	Δ	L	R	D	D	No	No	0.088[a] 0.178[b]	[45, 195, 198]
Hap V	N	R	R	D	E	Yes	Yes	0.202[a] 0.054[b]	[45, 198]
Hap VI	Δ	L	G	K	D	No	No	0.026[a] 0.0004[b]	[45, 198]
Hap VII	N	R	R	K	E	Yes	Yes	0.009[a] Not Detected[b]	[45, 198]

[a][198]

[b] [45]

Notably, A3H Hap II has some variability in its restriction ability which is dependent on alternative mRNA spliced forms [195]. A3H Hap II is primarily found in Africans/African Americans (~50%) and to a much lesser extent within other cultural populations (prevalence of approximately 0-8%) [45, 198]. Furthermore, an inverse correlation between the presence of Hap I and Hap II in different human populations has been established [202], where when A3H Hap II is common then A3H Hap I is rare (Africa) and when A3H Hap I is common then A3H Hap II is rare (Southeast Asia). It has been proposed that A3H evolved to become unstable due to a combination of the loss of an ancient pathogen and the ability of an ancestral A3H to induce mutagenesis of genomic DNA [36, 45]. Indeed, human A3H Hap I has been identified as an endogenous mutagen based on the findings that A3H Hap I has nuclear localization propensity, demonstrates enzymatic activity and is associated with the APOBEC mutation signature in breast tumours lacking A3B [202]. In accordance with the idea that A3H Hap I contributes to cancer mutagenesis, Zhu *et al*. [203] found that the T allele of rs139293, a SNP on A3H exon 2 that resulted in amino acid substitution from Arginine to Leucine at codon 18 and was predicted to have a deleterious effect on the structure and function of A3H, was significantly associated with reduced risk of lung cancer in Chinese population [203]. Therefore, it appears inactive forms of A3H may protect people from developing lung cancer.

A3H is the only common A3 enzyme with highly diversified antiviral activities based on sequence polymorphisms [197, 204] and appears to be in a category of its own in relation to other A3 enzymes regarding two other aspects. First, A3D, A3F, and A3G that also restrict HIV replication have two Z-type domains, whereas A3H has only one Z-domain (Fig. 1.3A) [37]. Phylogenic analyses have shown that the A3 members have three distinct Z-type domains (Z1, Z2, and Z3) and A3H is the only A3 enzyme with a Z3 (Fig. 1.3A) [37]. A3D and A3F have two Z2 domains and A3G has a Z1 (CTD) and Z2 (NTD) domain (Fig. 1.3A) [37]. Second, A3H is the only single Z-type domain A3 (others are A3A and A3C) that forms oligomers and multimers. Structural and biochemical studies have found that A3A and A3C are largely monomeric (>90%) in solution and do not multimerize in cells through an RNA intermediate [191, 205-209]. In contrast, A3H Hap II was found to multimerize in cells [191]. The A3H Hap II multimerization in cells was shown by fluorescence fluctuation spectroscopy and determined that multiple A3H Hap II molecules could closely associate on RNA, not that A3H Hap II oligomerized through a protein-protein interaction [191]. However, we have found using size

exclusion chromatography that A3H Hap II can form a dimer in solution in the absence of RNA or DNA [66, 210]. A3G and A3F also form oligomers in the absence of nucleic acid supporting the idea that A3 oligomerization ability facilitates the multimerization of A3 enzymes with RNA in cells [133, 140, 150]. It has been shown that similar to A3G and A3F, A3H Hap II interacts with cellular RNA and the NC portion of Gag to facilitate its encapsidation into HIV particles [198, 201]. Studies on A3H Hap II and Hap I have also shown that cytoplasmic localization correlates with restriction efficiency since mutation of A3H Hap I to make it cytoplasmic (G105R) increases its restriction capacity despite other amino acid differences from A3H Hap II [195, 211]. Additionally, virion mislocalization of certain A3H haplotypes may render them less active against HIV [212]. For example, despite nuclear localization of A3H Hap I, it can be encapsidated into HIV particles, but through an association with the matrix and capsid region of Gag, which leads to its primary localization outside the capsid [212]. These data suggest that both cellular and virion localization play a role in restriction efficiency.

Studies on the biochemical properties of A3H suggests A3H Hap II prefers to deaminate ssDNA at 5'TC sites, similar to A3F and A3D [210, 213, 214]. Through biochemical characterizations of A3H Hap II and Hap V, it was determined that A3H can utilize both 3-D jumping and intersegmental transfer and 1-D sliding during ssDNA scanning, and that the 3-D translocation mechanisms of jumping and intersegmental transfer are redundant [210]. A3H primarily dimerizes in solution through amino acid residues [44]RGY[46] on β2 strand and upon ssDNA binding it tetramerizes through amino acid residues [112]YYHW[115] on loop7 [210]. A3H is the only Z3-type deaminase [37] and appears to be the only A3 that oligomerizes in this manner. The C-terminal domain (Z1-type domain) of the double-domained A3G dimerizes through CTD loop7 (Fig. 1.5B) [138, 140] and the Z2-type single-domained A3C, A3F CTD and rhesus macaque A3G NTD dimerize through helix 6 [182, 215, 216].

1.3.5 A3C I188-mediated restriction of HIV-1

A3C has the lowest nucleotide diversity amongst all *A3* genes [204]. Although the common human A3C with a Serine (S) at amino acid position 188 has weak antiviral activity against HIV, A3C is polymorphic at this site and I188 is present at around 10% frequency in the African population [82, 204]. A3C I188 also demonstrated a 3 to 10-fold more potent antiviral activity as compared to the common human A3C [82, 216] and the difference in the anti-HIV activity of the A3C S188 and I188 was due to the enzymatic efficiencies of the two A3Cs. The stronger enzyme

specific activity and processivity of A3C I188 correlates with its increased dimerization [82, 216].

1.3.6 Restriction of HIV-1 by coordinately expressed A3 enzymes

Vif-deficient HIV was found to show replication defects when produced from cell lines such as CEM and H9, resulting in their classification as nonpermissive cell lines [217-222]. After many years of investigating the function of Vif and trying to understand the dichotomous phenomenon of permissive and nonpermissive cell lines for △Vif HIV, Sheehy *et al.* [25] found that the nonpermissive CEM cell line expressed A3G. Thereafter, many groups discovered that Vif was required to induce degradation of A3G to enable HIV replication [63, 64, 70, 223-225]. Later, it was realized by analyzing the mRNA expression levels of A3s using quantitative PCR in permissive (CEM-SS, SupT1) and non-permissive (CEM and H9) T cell lines that the classical nonpermissive CEM T cell line expressed not only A3G but also A3F and A3D, albeit with lower mRNA levels [116].

Because different A3s demonstrate unique mutation signatures (see above), evidence for multiple A3s mutating the same integrated proviral genome has been obtained when mutation sequence contexts indicative of A3-mediated deaminations have been identified [111, 168, 226-229]. It has been found that A3G can co-package into the same virion with A3F or A3H without competing for encapsidation and these A3s also act cooperatively to increase HIV inhibition by co-mutating the same genome in a single round of replication in both cell culture and in proviral DNA recovered from primary peripheral blood mononuclear cells (PBMCs) of chronically infected HIV-1 patients [117]. Ara *et al.* [230] investigated mechanistically how co-expression increases the mutagenic activities of A3G and A3F and they found that one molecule of A3G and three molecules of A3F can form hetero-oligomers in the absence of RNA and this A3F/G hetero-oligomer has distinct biochemical properties and functionalities from the individual enzymes [230]. A3F appears to increase the processive jumping ability of A3G when they form a hetero-oligomer and this enhances the ssDNA scanning abilities for the deamination motif, which serves as an indicator of A3's mutagenic potential [133, 230]. On the other hand, that A3G increases the amount of mutations being induced by A3F was due to A3G allowing A3F to inhibit reverse transcription more efficiently, which provides more time for A3F to search for and deaminate the target motif in (-)cDNA [230]. In regards to other A3 enzymes' abilities to cooperate during HIV infection, it has been found that A3D cannot co-mutate the genome along

with A3G efficiently because of its low virion encapsidation level [117], A3H activity can be increased synergistically in the presence of A3G [117], and that the combination of A3F and A3H activities is much weaker than the combination of A3G and A3F activities [229].

In primary CD4$^+$ T cells, A3 enzymes relevant to HIV restriction are expressed and further induced by mitogens, rather than interferon, indicative of their function in restricting retrotransposons [116, 231]. In contrast, in macrophages, monocytes and dendritic cells expression of select A3 enzymes is induced by interferon [116, 231]. Although A3 enzymes are not individually expressed in cells as shown in many laboratory experiments [80, 116], there is an advantage of individual expression of each A3. Individual expression of A3 enzymes enables mechanistic information to be obtained about enzyme function and mutational footprints to be established. However, A3s with perhaps a lesser restriction efficiency would not be expressed alone during an HIV infection suggesting that it may not matter *per se* which enzyme is most effective since they may each contribute cooperatively to HIV restriction. Gillick *et al*. [157] found that in primary human CD4$^+$ T cells the majority of proviral mutations were in a sequence context that indicated A3G-induced mutations are dominant (5'CC), but A3F- and A3D- induced mutations (5'TC context) were evident at 9-fold less frequency than the 5'CC context in △Vif HIV. This is in contrast to a study reported by Ooms *et al*. [100] that used PBMCs to examine the hypermutation of HIV in the absence or presence of A3H Hap II. In the absence of A3H Hap II it was found using a deep sequencing approach that there was approximately an equal number of mutations originating in 5'CC and 5'TC contexts, suggesting that A3F and A3D cooperate to induce a equivalent number of mutations to A3G [100], in agreement with results from a CEM2n T cell line [80]. In the presence of Vif that could induce degradation of all A3s except A3H Hap II, there was a large number of mutations in the 5'TC context demonstrating that A3H when present in a stable form is highly active against HIV.

Although the use of spreading infections in primary cells or T cell lines support the idea that A3s cooperate, there still may be a question of whether they induce HIV evolution. It has been proposed that if there is an insufficient amount of A3-induced hypermutation this may benefit HIV and contribute to sequence variation by induction of sublethal levels of mutagenesis which results in HIV evolution [74, 109, 118, 124]. There is evidence that A3G and A3F hotspots are enriched in immunogenic CTL epitopes and that HIV may utilize A3s to induce immune escape [109]. In addition, A3G may be able to induce resistance to the RT inhibitor lamivudine (3TC)

because its deamination motif overlaps with a codon for Met and results in a M146I mutation in the *pol* gene [74, 118]. However, the frequency of this mutation being induced by A3G versus RT activity has been questioned [126]. It is also not known if A3F/A3D/A3H Hap II induce this evolution any more than A3G, due to differences in inactivation potential derived from their sequence specificities [91, 133, 172, 206] and if this impacts disease progression in infected individuals. On the other hand, Vif has been shown to adapt within HIV infected individuals and to be less effective in inducing A3 degradation [73, 125, 232]. It is thought that HIV can utilize Vif as a mutational rheostat in times of viral stress by allowing low amounts of A3s into viral particles to induce sublethal mutagenesis [73, 125]. These types of studies have raised the idea that perhaps inducing hypomutation or shutting off A3 enzymes may benefit HIV infected individuals [233, 234].

1.4 HIV-1 Viral infectivity factor (Vif)

1.4.1. General properties of Vif

The main function of Vif was not fully understood for the first decade of HIV research. It was not known why, but the presence of Vif made some cell lines permissive for producing HIV particles capable of undergoing another round of infection [235, 236]. Nonpermissive cells allowed a △Vif HIV to produce virus particles, but they were rendered noninfectious upon infection of fresh cells. Two labs discovered that Vif repressed a host factor [221, 222]. It was later identified by subtractive hybridization that A3G (originally called CEM15) was the host factor that was highly packaged into virions in the absence of Vif and blocked infection in the next target cell (Fig. 1.4A) [25]. Although this is clearly a primary role for HIV infectivity, Vif was also shown to influence HIV particle morphology and this may relate to its potential role as a nucleic acid chaperone [220, 237-239]. Although Vif is expressed by all HIV strains, its sequence varies considerably amongst different HIV isolates [240].

1.4.2 Vif as an E3 ubiquitin ligase substrate receptor

Canonically, CBFβ functions as a transcription cofactor to the RUNX family of transcription factors to regulate the expression of genes involved in T lymphocytes and bone cell development [241]. In 2012, it was discovered that Vif interacts with CBFβ for stability in cells [68, 69]. The interaction is mediated through Vif amino acids [84]GxSIEW[89] and [102]LADQLI[107] (Fig. 1.7, A and B) [242, 243]. The Vif/CBFβ complex is also required for *in vitro* stability of Vif

A

A3D

A3F

A3H

M // WQxDRMR // WKSLVK // FYRHHY // H // YWxL // TGERxW // GxSIEW // LADQLI // Hx$_2$YFxCFx$_4$IRx$_2$LxGx$_6$CxYx$_3$H // SLQYLA // PPLPx$_4$L // H

A3G

A3C

Cul5-box/Zinc finger region

PPLP Motif/ Dimerization Region

CBFβ

BC-box

B

Vif

A3F/C/D
^{14}WQxDRMR^{17}

CBFβ ^{84}GxSIEW^{89}
^{102}LADQLI^{107}

EloC
^{144}SLQYLA^{149}

^{108}H

α/β-domain

A3H
^{48}H

α-domain

^{114}C

^{139}H

A3G
^{40}YRHHY^{44}

^{133}C

A3H
^{39}F

Cul5
^{120}IRxxL^{124}

C

Vif/CBFβ/EloB/C/nCul5

EloB

Vif

CBFβ

EloC

nCul5

34
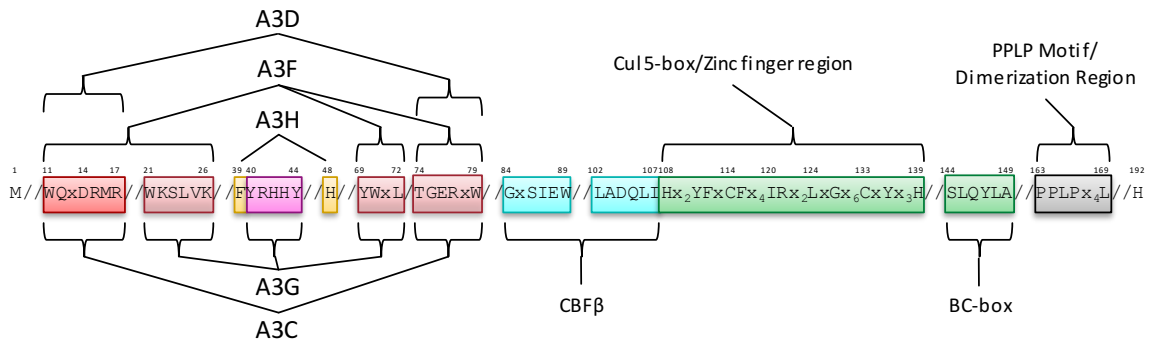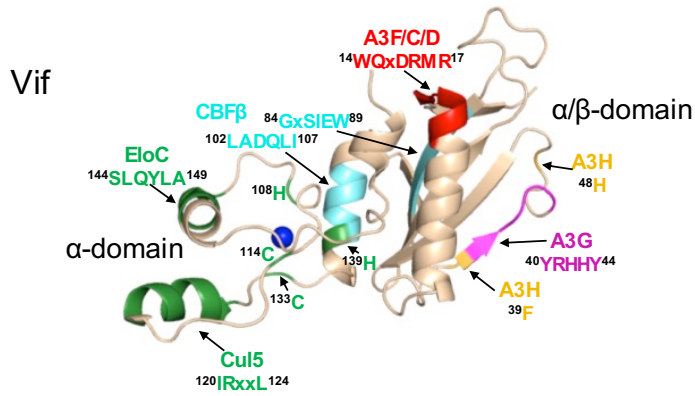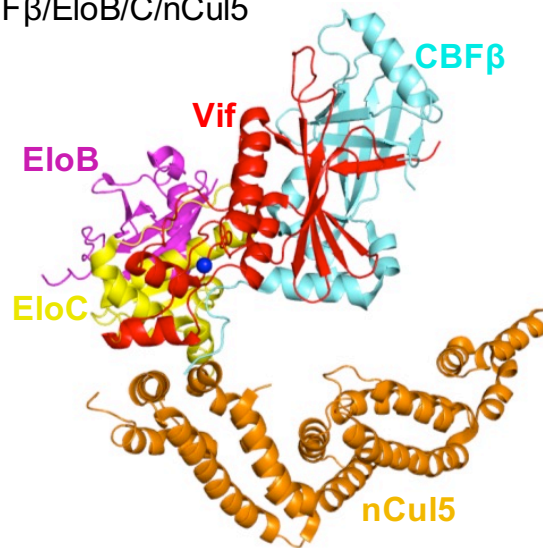
Figure 1.7 Structure of Vif and host interacting partners. (A) Domain organization of Vif. Vif uses specific motifs to interact with A3G (magenta, ^{40}YRHHY^{44}), A3F/A3C/A3D (red, ^{11}WQxDRMR^{17} and ^{74}TGERxW^{79}), and A3H (orange, ^{39}F and ^{48}H). In conjunction with these specific motifs, there are shared interaction motifs for A3F and A3G with Vif (pink, ^{21}WKSLVK^{26} and ^{69}YWxL^{72}). CBFβ interacts with Vif through two adjacent motifs (cyan, ^{84}GxSIEW^{89} and ^{102}LADQLI^{107}). The Zinc finger region (green, amino acids 108-139) coordinates the Zinc through an ^{108}H^{114}C^{133}C^{139}H motif and stabilizes Vif structure, which indirectly enables an interaction with Cullin 5 (Cul5). Direct interaction of Vif with Cul5 is through amino acids ^{120}IRxxL^{124}. The BC box mediates an interaction with Elongin C (green ^{144}SLQYLA^{149}). Vif oligomerizes through a PPLP motif (grey, ^{163}PPLPx$_4$L^{169}). Slanted lines are used to indicate intervening amino acids between the domains. (B) The crystal structure of Vif (wheat) shows that it has two domains on either side of a bound Zinc (blue). The N-terminal α/β-domain consists of a five stranded β-sheet, a discontinuous β-strand and three α-helices. The α/β-domain contains the binding interface for CBFβ (cyan, ^{102}LADQLI^{107}, ^{84}GxSIEW^{89}) and A3 enzymes. The ^{11}WQxDRMR^{17} motif (red) is used to interact with A3F, A3C, and A3D, the ^{40}YRHHY^{44} motif (magenta) is used to interact with A3G, and residues ^{39}F and ^{48}H (orange) are used to interact with A3H. The α-domain contains two alpha helices that mediate two separate interactions with EloC (green, ^{144}SLQYLA^{149}) and Cul5 (green, ^{120}IRxxL^{124}). (C) Structure of HIV Vif (red) in complex with CBFβ (cyan), Elongin C (EloC, yellow), and the N-terminal domain of Cullin 5 (nCul5, orange). Elongin B (EloB, magenta) dimerizes with EloC. Figures were made using PyMOL (The PyMOL Molecular Graphics System, Version 1.5.05, Shrödinger, LLC.). Figure reprinted from Feng *et al.*, 2014 [66].

and enables recombinant expression of Vif in a largely soluble form in *E. coli* that can be purified for biochemical studies [244]. In contrast, Vif alone expressed in *E. coli* accumulates in inclusion bodies and must be purified under denaturing conditions [245]. To suppress A3 action Vif interacts directly with A3 enzymes and mimics the human protein suppressor of cytokine signaling-2 (SOCS2) to become the substrate recognition subunit of a Cullin 5 (Cul5) E3 ubiquitin ligase complex (Fig. 1.4B).

Vif interacts with host proteins Elongin C, an obligate heterodimer with Elongin B (EloB/C) and Cul5 [70, 246-251]. The interaction of Vif with EloB/C increases the stability of Vif in cells and *in vitro* and promotes recruitment of CBFβ [252]. The interaction of Vif with EloC is mediated through a SLQ motif in Vif termed the Elongin B/C (BC) box (Fig. 1.7, A and B, ^{144}SLQYLA^{149}) [70, 91], similar to human SOCS proteins [253, 254]. Distinct from human proteins is that Vif does not have the highly conserved Cys in the BC box and instead has a ^{149}A [253-255]. The data with Vif suggests that it is the short side chain of the amino acid at position 149 rather than the Cys that is required for the interaction with EloC [70, 250]. Vif also does not contain a canonical Cul5 box [248, 249]. In the search for the presence of conserved Cys in the BC box, two other Cys (^{114}C, ^{133}C) were identified in Vif upstream of the BC box and were found to be involved in binding with Cul5 [70, 247]. These Cys were found to be part of a novel Zinc binding HCCH motif (Fig. 1.7, A and B, ^{108}Hx$_2$YFxCFx$_4$IRx$_2$LxGx$_6$CxYx$_3$H^{139}). The Zinc coordination in the HCCH was predicted to stabilize a small domain of Vif and indirectly support Cul5 binding [248]. The primary Vif amino acids that contact Cul5 are hydrophobic amino acids at positions 120 and 123-124 in a helix that is adjacent to the HCCH residues (Fig. 1.7, A and B, ^{120}IRxxL^{124}) [72, 256]. The Vif/CBFβ/EloB/C heterotetramer undergoes a conformational change that promotes binding to Cul5, suggesting that there is a prescribed order in the assembly of the E3 CRL5 ligase complex [257]. Accordingly, Cul5 binds less well to EloB/C in the absence of Vif/CBFβ [256].

A recent structural study of Vif bound to CBFβ/EloB/C/Cul5 shows that Vif has an overall elongated cone structure and contains two domains with a Zinc binding domain in the center of the two domains (Fig. 1.7C) [256]. CBFβ binds the N-terminal α/β-domain and the EloC and Cul5 bind the C-terminal α-domain of Vif (Fig. 1.7C) [256]. Both EloC and Cul5 interact with Vif through hydrophobic interfaces on distinct α-helices [249, 256]. The crystal structure also

emphasizes the stability that CBFβ imparts to Vif since they have a total interaction surface area of 4,797 A^2 and form an antiparallel β-sheet with a β-strand from each protein [256]. The side of CBFβ that is bound by Vif is the same side that the human CBFβ binding partner, the RUNX1 transcription factor, binds to, suggesting a mutually exclusive binding [256, 258], although other reports show CBFβ can bind Vif and RUNX1 on distinct surfaces [69, 259, 260]. Functionally, Vif appears to exclude CBFβ from binding RUNX1 because overexpression of Vif can alter the transcription profile of nearly 100 genes that are known to be regulated by RUNX1 in Jurkat T cells and suggests that Vif may have multimodal effects in HIV infected cells [258]. In support of this, Anderson *et al*. [261] found that CBFβ interaction with RUNX protein is required for *A3* gene expression in CD4$^+$ T cells, indicating that by hijacking CBFβ, Vif is able to suppress A3 activity both post-translationally through degradation and transcriptionally by precluding RUNX/ CBFβ complex formation. CBFβ is only required by primate-specific pathogens, such as HIV and SIV, for Vif to function [259] and CBFβ is dispensable for the activity of non-primate lentivral Vif proteins. Lentiviruses that infect cats (feline immunodeficiency virus), cattle (bovine immunodeficiency virus) or sheep (maedi-visna virus, MVV) appear to lack an analogous cofactor or to have adapted to use a different cellular cofactor. For example, Cyclophilin A (CypA) [262] is used for MVV to stabilize Vif [262, 263]. Because mammalian hosts of these viruses have simpler A3 repertoires [37], it has been proposed that adaptation to hijack CBFβ and overcome A3s at both the transcriptional and post-translational level is a 'gain-of-function' adaptation that Vif adopted in order to conquer the zoonotic barrier created by the expanded A3 repertoire of primates during a cross-species virus transmission [261, 263, 264].

CBFβ interacts with a hydrophobic region of the Vif α/β-domain, but the rest of the exposed α/β-domain surface is highly positively charged and is thought to mediate electrostatic interactions with A3 enzymes [256, 265]. To target A3s for proteasomal degradation, Cul5 interacts with RING box protein 2 (Rbx2) [68] and this results in the assembly of a hexameric complex (Fig. 1.4B). Further, a ubiquitin conjugating E2 enzyme interacts with the hexamer through Rbx2 and causes polyubiquitination of the A3 enzyme, on multiple lysine residues through a ^{48}K linkage signaling it for degradation through the proteasome pathway (Fig. 1.4B). Current data for A3G and A3F suggest that the Lys residues that become conjugated to ubiquitin are random [266].

1.4.3 Vif: amino acids that interact with A3s

Alanine scanning mutagenesis of Vif or comparison of different Vif variants from HIV subtypes has enabled the identification of three distinct regions of Vif that interact with A3G, A3F/A3D/A3C, or A3H [175, 240, 267]. Vif interacts with A3G on two widely distributed motifs. Two positively charged motifs on Vif, ^{21}WxSLVK^{26} and ^{40}YRHHY^{44} are involved in the interaction with A3G (Fig. 1.7, A and B) [268-272]. In addition, the ^{69}YWxL^{72} motif is a region of Vif that interacts with both A3G and A3F (Fig. 1.7A) [273, 274]. However, for both A3G and A3F mutation of the ^{40}YRHHY^{44} and ^{14}DRMR^{17} motifs to all alanines are necessary and sufficient to block Vif-induced A3G and A3F degradation, respectively, suggesting the other domains provide a secondary stabilizing interaction [269]. Of note, A3C and A3D share a common binding site on Vif as A3F, with ^{14}DRMR^{17} shown to be of importance (Fig. 1.7, A and B) [205, 274, 275]. Vif interacts with A3H through another unique site that involves amino acids ^{39}F and ^{48}H (Fig. 1.7, A and B) [240, 276].

The Vif amino acids that interact with A3H are not highly conserved among HIV subtypes, in contrast to the motifs of Vif that interact with A3G and A3F. It has been suggested that since HIV rarely encounters a host with A3H Hap II that is effective against HIV, there has been evolutionary drift of Vif to not maintain an interaction site with A3H [100]. As a result, A3H is differentially sensitive to Vif variants. For example, A3H Hap II is not sensitive to HIV NL4-3 Vif (^{39}F, ^{48}N), but is sensitive to HIV LAI Vif (^{39}F, ^{48}H) [276]. The inability of some Vif variants to induce degradation of A3H Hap II enabled Ooms *et al*. [100] to test whether A3H could act as an infection barrier to HIV. They found that Vif will adapt in infected individuals to induce degradation of A3H Hap II. Importantly, this evolution of Vif affects only A3H and Vif maintains the ability to induce degradation of A3G and A3F [100], confirming that Vif indeed uses three distinct interfaces to interact with A3 enzymes and supporting the idea that multiple A3 enzymes coordinately exert a restriction pressure on HIV. Importantly, treatment naïve HIV-infected individuals at the early or primary infection stage that had at least one active A3H allele (Hap II) had higher levels of mutations in proviral genomes in a 5′TC context, lower viral loads and higher CD4$^+$ T cell counts [100]. Gourraud *et al*. [277] similarly reported that early stage untreated HIV infected individuals that were homozygous for a stable A3H allele demonstrated lower HIV RNA over time, but this did not correlate with increased hypermutation of HIV proviral genomes. This difference in mutational load between these reports is likely due to the

different sequencing strategies used in the two studies [100, 277]. These data are similar to clinical data obtained with A3G and A3F that demonstrate in a number of cohorts (but not all), A3G or A3F mRNA expression or hypermutation levels correlate with high CD4$^+$ T cell counts and low viremia (reviewed in [115]). Based on these studies [100, 277], Refsland *et al*. [232] took a step further and reported that the ability of A3H to restrict HIV depends on both polymorphisms in host *A3H* genes and natural variations in the Vif proteins of the infecting virus. They found that due to the selection pressure A3H exerted on HIV, Vif variants that are most capable of neutralizing A3H also dominate in areas where stable A3H haplotypes prevail and *vice versa*. Hence, Vif appears to be able to adapt to different A3H haplotypes in a population. This means stable A3H may act as a *bona fide* transmission barrier and lower the rate of virus transmission during early infection when a HIV strain from a patient with an unstable haplotype (thus carrying a Vif variant unable to neutralize A3H) infect a person with a stable haplotype, as Vif needs adaptation to re-gain its ability to neutralize A3H [232].

1.4.3.1 Vif-A3G interaction

The Vif-A3G interaction was the first Vif-A3 interaction to be studied and it established that Vif inhibits the antiviral activity of A3 enzymes in a species-specific manner [67, 278-282]. This means that HIV Vif cannot neutralize A3G from African green monkey (AGM), and AGM SIV Vif cannot neutralize A3G from human and this has been recognized as a cross species infection barrier [67, 278-282]. Initially, to identify the residues HIV Vif uses to interact with human A3G the human A3G amino acids were replaced with those of AGM or rhesus macaque A3G. Mutation of human A3G ^{128}D to ^{128}K as found in AGM and rhesus macaque A3G abrogated the interaction of HIV Vif with human A3G and its ability to induce degradation of human A3G [278-280]. However, when the ^{128}D was mutated to ^{128}A, HIV Vif could still interact with and degrade human A3G, demonstrating that the charged interface was more important than the amino acid identity [280]. Since mutation of solely ^{128}D to ^{128}K can abrogate the interaction between A3G and Vif in co-immunoprecipitation studies it is clearly a determining residue; however, Vif mediated degradation can be influenced by mutation of A3G ^{129}P and ^{130}D and Vif also interacts with A3G on surrounding motifs such as helix 6 (Fig. 1.5A) [183, 267, 283]. The loop 7 and helix 6 region contain more positively charged and neutral amino acids than negatively charged amino acids which may explain why ^{128}D is such an important contact point for the positively charged Vif, despite a larger A3G interface predicted from biochemical studies.

The A3G-Vif interface identified by Letko *et al.* [284] through functional assays and molecular docking confirmed that A3G loop7 (also known as β4-α4 loop) indeed serves as a core binding site of Vif and amino acid residues 125, 128 and 130 fit into a well-defined positively charged pocket on the surface of Vif. Other A3G loop7 residues such as 124, 127, 129 and 131 interact with Vif residues lining the pocket [284]. Importantly, although primate-specific amino acid polymorphisms exist in A3G loop 7, this structure is conserved across different primate species and it provides an explanation for why species-specific variation in loop7 determines primate-species tropism [284]. Nevertheless, these studies with A3G established the principle that a lack of Vif-induced degradation correlates with a lack of an interaction between the A3 and Vif.

1.4.3.2 Vif-A3F/A3D interaction

Unlike A3G where the Vif binding and the enzyme catalytic activities reside within the NTD and CTD, respectively, both these activities reside in the same domain, CTD, for A3F and A3D (Fig. 1.5, D and E) [205, 285]. A3F has been studied more extensively than A3D in this regard and will be discussed here. In contrast to A3G, there was no specific single amino acid determinant being identified for A3F that clearly mediated both the primary interaction with Vif and was a determinant for Vif-mediated degradation. Rather, different groups identified different amino acids in A3F that altered its susceptibility to Vif [179, 205, 269, 285, 286]. For example, [289]E has been reported to be critical for A3F sensitivity to Vif [285, 287]. Albin *et al.* [288] and Land *et al.* [286] identified another residue, [324]E, as the key determinant of A3F to Vif-mediated degradation. Based on their molecular dynamic (MD) simulation-optimized model, Richards *et al.* [287] found that the A3F-Vif interaction is dominated by a stable electrostatic interaction between A3F [289]E and Vif [15]R and multiple A3F amino acid residues identified through prior mutagenesis studies [179, 205, 269, 285, 286] all located on the solvent-exposed surfaces of the helix 2,3, and 4 and can be re-grouped into contacts that either directly or indirectly involve in the actual A3F-Vif interaction. In agreement with this study [287], a combination of mutagenesis, structural modeling and a crystal structure of A3C that shares the same Vif binding interface with A3F and A3D, identified a novel type of A3 and Vif interaction [205, 285], where Vif interacted with a negatively charged surface of A3F/A3D/A3C that spanned helix 2, 3 and 4 and β-strand 4 (Fig. 1.5, D and E) [205, 265], rather than Vif interacting with a loop as in the case of A3G (Fig. 1.5A, loop7). This negatively charged surface supports the hypothesis that it is primarily electrostatic interactions that mediate the A3 and Vif interaction and provides an

explanation for why the A3F and Vif interaction may be more difficult to disrupt than the primarily neutral surface present in A3G. Whether this would mediate a tighter interaction of Vif with A3F than A3G is not known since there are no quantitative data available for both A3G and A3F using the same experimental conditions. Studies with A3F have shown that a lack of Vif-induced degradation does not necessarily correlate with a lack of a Vif-A3F interaction, suggesting that the binding orientation or other factors contribute to successful Vif-mediated degradation rather than only the presence of an interaction [288].

1.4.3.3 Vif-A3H interaction

A3H sensitivity to Vif is haplotype dependent [45, 78, 195, 211, 240, 289]. The A3H Hap I is not sensitive to HIV LAI Vif-mediated degradation whereas A3H Hap II is sensitive to HIV LAI Vif-mediated degradation [195, 211, 276, 290]. The A3H haplotype polymorphisms only occur at three locations (amino acids 105, 121 and 178) (Table 1). A3H Hap I encodes GKE at these three positions and A3H Hap II encodes RDD at these positions. It was shown that at position 105 the Arg is required for stable expression in cells and that the 178 position had little effect on Vif-mediated degradation [45, 195, 211]. Therefore, a single amino acid homologous to A3G [128]D at position 121 in A3H Hap II was determined to control sensitivity to Vif-mediated degradation (Fig. 1.5F) [211, 290]. An A3H Hap II mutant with a [121]K is not sensitive to Vif mediated degradation and does not interact with Vif [290]. However, a recent study that identified the A3H-Vif interface through structure docking using the A3G CTD and Vif crystal structure as models [256] [291] suggests instead of a single amino acid [121]D, the Vif-A3H interaction sites can be mapped to α-helix 3 (involves amino acid position 86, 87, 90 and 94) and α-helix 4 (involves amino acid position 121, 125, and 129) in A3H and β2 to β5 strand in Vif [292]. Two Vif residues at amino acid position 39 and 48 have been previously identified to be important for A3H interaction [100, 240, 276] and have been found to indeed locate at the A3H-Vif interface. Specifically, Vif F39 interacts with A3H amino acids at position 121, 124 and 125 and Vif H48 interacts with A3H amino acid at position 90. Interestingly, they also found Vif amino acid 40-44 on β2 sheet is important for interaction with both A3G and A3H, which was thought to be exclusively used by A3G [175, 229, 284], although the modeling predicts A3G and A3H interact with Vif in this area on different sides [292]. Notwithstanding, from the A3H structural model (Fig. 1.5F) it appears that the [121]D of A3H is not located on loop 7 as in A3G, but is on helix 4 and on a different face of the molecule (Fig. 1.5, A and F). Yet, similar to A3G,

41

the region surrounding [121]D in A3H contains mainly neutral or positively charged residues, in contrast to the negatively charged interface that Vif uses to interact with A3F and A3D [265].

1.4.4 Degradation-independent inhibition of A3G

Although the primary mechanism by which Vif inhibits A3 enzymes is through inducing their proteasomal degradation, there have been other ways identified in which Vif can inhibit A3G encapsidation or function through a degradation-independent route. Vif may not be able to completely induce degradation of A3G in the virus producing cell and these degradation-independent mechanisms may be another line of defense against A3G virion encapsidation. In particular, Vif can become the target of A3-mediated hypermutation [73, 126], which may result in a Vif unable to interact with the E3 CRL5 ligase complex, but still able to inhibit A3G through a degradation-independent mechanism. It is not known if Vif can act in this manner for other A3 enzymes

1.4.4.1 Vif decreases translation of A3G mRNA

Vif can decrease A3G mRNA translation in order to lower the steady state levels of A3G through a Vif and A3G mRNA interaction, but the exact mechanism is not understood [223, 225, 293]. It is known that Vif can decrease the mRNA levels of A3G by 15-40% and this requires that Vif interact with the 5'UTR of the A3G mRNA [225, 293] and stem-loop 2 and 3 within the 5'UTR of A3G mRNA are required for the translation inhibition by Vif [294]. Importantly, Guerrero *et al*. [294] demonstrated that the translational repression of A3G and proteasomal degradation of A3G induced by Vif can independently contribute to the overall decrease of A3G intracellular and intravirion level and that the importance of the translational control of A3G by Vif was overlooked because previous studies have used A3G expression vectors that lack an authentic A3G 5'UTR [294]. The ability to translationally inhibit A3G is well conserved amongst different Vif variants and interestingly Vif variants from HIV lab strains LAI and NL4-3 showed weaker inhibition of A3G translation as compared to variants of clinical origin, such as A2 and D1[294]. Nevertheless, since Vif has been shown in an immunofluorescence study to co-localize with A3 enzymes and P-bodies [295], it is possible that Vif shuttles A3G mRNA to P-bodies to delay or prevent its translation.

1.4.4.2 Vif inhibits virion encapsidation of A3G

Studies by Goila-Gaur *et al*. [296] have shown that A3G synthesized *in vitro* using a rabbit reticulocyte lysate translation system would become immunoprecipitation and packaging incompetent in the presence of Vif. Vif was not associated with these high molecular mass A3G

forms, but was required for their formation [296]. Although A3G regularly forms high molecular mass complexes in cells, which are less likely to be packaged into virions, Vif can induce an even higher molecular weight form of A3G [135, 296]. Moreover, studies with an A3G C97A mutant that is resistance to Vif-mediated degradation suggested that Vif-mediated degradation and inhibition of packaging are two distinct properties of Vif since the A3G C97A mutant was encapsidated less well in the presence of Vif [297]. A molecular mechanism for this effect has not been described.

1.4.4.3 Vif inhibits deamination of deoxycytidine by virion-encapsidated A3G

A3 enzymes are mainly studied for restriction of HIV$\triangle vif$ in order to observe restriction in single-cycle infectivity assays, but in infected individuals A3 enzymes must contend with Vif. Despite multiple mechanisms that Vif uses to block A3G, it has been shown that A3G is encapsidated in the presence of Vif, albeit in lower amounts [93]. Further, per molecule of A3G there is less deamination activity (Fig. 1.4E) [298]. This decrease in A3G deamination activity occurs even when A3G and Vif are coexpressed in *E. coli* and mutations are detected with a Rifampin reversion assay or *in vitro* with purified A3G and Vif, demonstrating that other viral components are not required for the inhibition to take place [183, 298, 299]. Enzymatic studies have shown Vif can cause a decrease in the specific activity of A3G and this is due to a combination of Vif competitively binding to the ssDNA substrate and Vif binding directly to A3G [183, 298]. These are separable functions of Vif since ssDNA-binding studies of Vif-A3G complex in comparison to each of the component binding ssDNA alone supported the hypothesis that Vif bound to A3G is unable to bind ssDNA with high affinity [183]. Another consequence of Vif binding to A3G is that it disrupts how A3G scans ssDNA in search of cytosine to deaminate [183]. Vif interacts with the A3G NTD on loop7, which is required for processive jumping movements (Fig. 1.3A) [139]. In a study that used two Vif variants to examine the effect of Vif/CBFβ on the deamination activity of A3G it was found that HIV Vif HXB2 inhibited A3G jumping movements, consistent with an interaction of Vif on loop7 [183]. In contrast, HIV NL4-3 Vif inhibited A3G sliding, which is mediated by helix6, providing functional evidence that beyond the key loop 7 contact residues, ^{128}DPD^{130}, Vif variants can interact with different regions of A3G [183]. This appears to have no functional consequences for A3G-mediated degradation [240], but provides insights on how variable the Vif variants can be in their extended binding sites on A3 enzymes. This may affect strategies that aim to use

small molecule inhibitors of the Vif-A3 interaction as an HIV therapy. Altogether, the data suggest the mechanism by which Vif inhibits A3G deamination activity in virions is by altering the scanning mechanism used to find cytosine on ssNDA. Inhibition of A3G deamination activity by Vif is likely to result in sublethal mutagenesis of HIV and could contribute to the generation of viral quasispecies and HIV evolution [124, 183].

1.4.4.4 Development of small-molecule inhibitors for A3-based HIV-1 therapeutics

The Vif-A3 molecular interface between the host and pathogen creates the possibility of developing novel therapeutics [115, 300] and high-throughput screening approaches for small-molecule inhibitors have delivered positive results. There are strategies to induce either A3G-mediated viral hypermutations by disrupting the Vif-A3G interaction [301-306] or viral hypomutation by blocking A3G catalytic activity [307, 308]. For the "therapy by hypermutation" strategy, the rationale is to find small-molecule inhibitors that antagonize Vif function and increase the cellular level of A3G available for virus restriction. A few candidate molecules that recover A3G expression levels and enable HIV restriction in the presence of Vif have been discovered [301-306, 309], although there are little biochemical data to understand the mechanism of action. For the "therapy by hypomutation" strategy, small-molecule inhibitors have been designed that target a key residue in A3G (C321) that inhibits its catalytic activity [307, 308]. It is thought that decreasing the viral quasispecies that may arise due to A3-mediated mutagenesis can assist in immune clearance of the virus and decrease resistance to antivirals [233, 234]

Inhibitors targeted to Vif may only be successful if administered in a cocktail to cycle their use and prevent the development of drug resistance, a long-standing therapy regimen for HIV drugs [310]. One strategy to avoid selection of resistant Vif is to utilize the anti-HIV potential of each A3 enzyme and design inhibitors that bind different regions of Vif, based on the unique interactions that Vif has with A3G, A3F/D and A3H (Fig. 1.7A). However, more study is required to determine if all A3 enzymes function equally well as individual restriction factors otherwise, the strategy may need to involve inhibiting degradation of all A3 enzymes together to enable a strong restriction pressure on HIV. Development of inhibitors that target the A3 enzymes may be a problematic route when considering A3G since Vif interacts with A3G near the amino acid residues needed for virion incorporation, oligomerization and processivity (Fig. 1.3A) [138, 183, 267]. As a result, the inhibitor molecule may decrease A3G anti-HIV activity. It

is unknown whether the activity of A3D, A3H, or A3F would be affected by this type of strategy. Furthermore, it has been shown in the case of A3G, that HIV can overcome the restriction pressure of A3G by acquiring mutations in genetic sequences other than Vif in order to indirectly avoid A3G encapsidation [180, 181].

If Vif were unable to interact with E3 CRL5 ligase complex components, the accelerated degradation of A3G would be blocked. This strategy has been raised as a potential option [251, 300, 311] and current structural data on the EloB/C-, Cul5-, and CBFβ- Vif interfaces could facilitate development of inhibitors [250, 256]. However, the consequence of targeting the host proteins with small molecules remains unknown. In addition, this approach has potential drawbacks since Vif may remain bound to the A3 enzymes. For A3G, this has been shown to lead to a decrease in mutagenic activity [183, 298]. *In vitro*, Vif/CBFβ does not affect A3H deamination activity [312]. There are no published studies investigating whether Vif would affect the mutagenic activity of other A3s. If Vif were unable to interact with CBFβ it would become unstable in the host cells and degradation of A3 enzymes would be circumvented [68, 69]. However, targeting a small molecule to CBFβ may be problematic if this prevents CBFβ from functioning as the transcription cofactor for RUNX proteins. Although some reports show that Vif and RUNX1 interact with CBFβ on distinct surfaces [69, 259, 260], Kim *et al*. [258] demonstrated that Vif recruitment of CBFβ alters the transcriptional profile of the cell by preventing RUNX1 and CBFβ association, suggesting Vif and RUNX1 may share a mutually exclusive binding surface on CBFβ [256, 258].

1.5. A3s-mediated inhibition of endogenous retroelements

Endogenous retroelements and their mobility have played a critical role in shaping the human genome during speciation and creating genome diversity [313, 314]. However, the excessive mobility of retroelement interferes with genome integrity and thus uncontrolled transposition events need to be suppressed [34, 315]. A3 enzymes have been defending the genomic integrity of their host cells long before HIV was transmitted to the human population, based on the evidence of a long history of positive selection recovered on the A3 locus [316-318]. It has been suggested that the primitive function of A3s is to restrict the mobilization of mammalian endogenous retrotransposons and the selection pressure imposed by these retroelements has also helped shaping the complex locus structure of the present day A3s and contributed to the diverse functions of A3 family members [35, 319].

A

ENV RNA

5'LTR Gag Pol (Env) 3'LTR

Integrated ERV

3'
5'

VLP

Extrachromasomally-primed
reverse transcription

B

L1 RNA

5'UTR ORF1 ORF2 3'UTR

Integrated L1

3'
5'

ORF2   ORF1

L1 RNP

target primed
reverse transcription

3'
5'

L1 cDNA
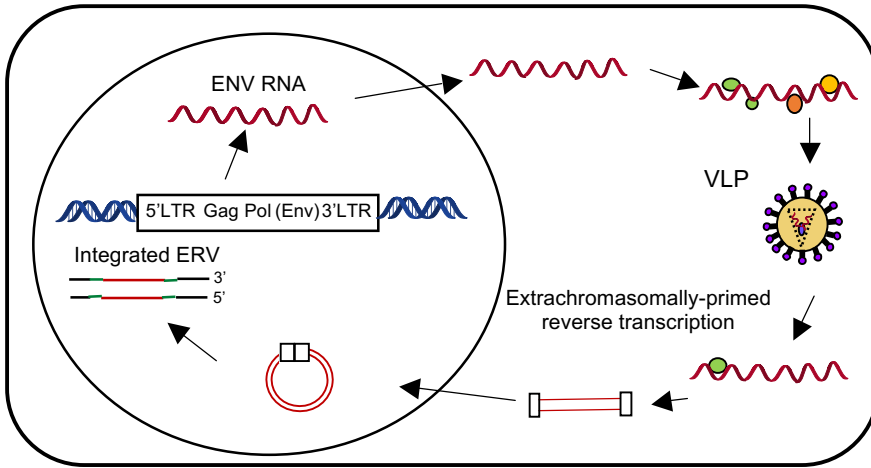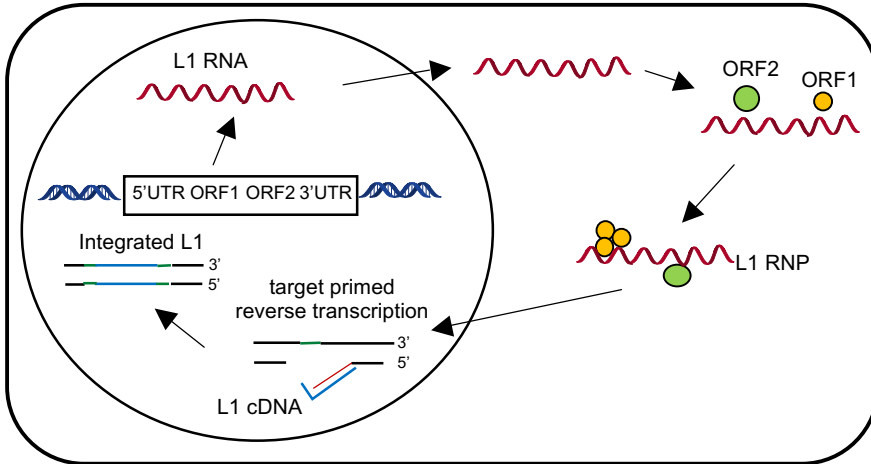
Figure 1.8 Schematic overview of the retrotransposition cycle of retroelements. (A) The retrotransposition cycle of the LTR-retrotransposons (also known as the endogenous retroviruses, ERVs). Individual ERVs comprise two long terminal repeats (LTRs), separated by approximately 5-10 kb of sequence encoding the canonical retroviral *gag*, *pol* and *env* genes (although *env* is usually defective). The life cycle of ERVs include the formation of virus-like particles (VLPs) that remain intracellular. Reverse transcription of ERVs occurs in the cytoplasm in a process called "extrachromosomally-primed reverse transcription (EPRT)". A structure that closely resembles the pre-integration complex (PIC) of exogenous retroviruses will be formed after EPRT, followed by an integrase enzyme (vertical box)-assisted integration event in the nucleus. The inserted ERV sequence is shown in red and the flanked target sites are shown in green. (B) The retrotransposition cycle of the non-LTR retrotransposons. A full-length non-LTR retrotransposon, represented by a L1 here, contains two open reading frames (ORFs), a 5'untranslated region (UTR) with internal promoter sequences and a 3'UTR with polyadenylation signals. ORF1 (yellow) encodes a nucleic acid binding protein with nucleic acid chaperone activity; ORF2 (green) encodes proteins with endonuclease and reverse transcriptase activity. After protein translation, ORF1p protein (ORF1p) and ORF2p preferentially act *in cis* and bind to their encoding mRNA, giving rise to the L1 ribonuclear protein (RNP) complex, an important L1 retrotransposition intermediate. The RNP complex then shuttles to nucleus for L1 replication and integration event. L1 propagates using a "copy and paste" mechanism and its DNA synthesis is based on a process named "target-primed reverse transcription (TPRT)", during which an ORF2p nicks the target DNA and exposes 3'OH, followed by polymerization of cDNA (blue) using L1 RNA (red) as a template. After integration, the entire L1 element is usually flanked by target site duplications (TSD), with the duplicated target sites shown in green.

Endogenous retroelements, including the LTR and non-LTR retrotransposons, represent a class of the mobile genetic elements that replicate through an RNA intermediate (for review, see [320, 321]) (Fig. 1.8). LTR retrotransposons are also called endogenous retroviruses (ERVs) and they account for ~10% of the mammalian genome [322]. The replication of LTR retrotransposons is also called extrachromosomally-primed retrotransposition [323] and the reverse transcription process occurs exclusively in the cytoplasm of the infected cells (Fig.1.8A). LTR retrotransposons structurally resemble exogenous retroviruses such as HIV, except they usually lack a functional *envelope* (*env*) gene, which means they can only propagate by integrating into the genome at a new site within their host cell of origin and thus are transmitted vertically. Non-LTR retrotransposons, on the other hand, are also called target-primed retrotransposons (Fig. 1.8B) [323] and the reverse transcription predominantly takes place in the nucleus using a process named "target primed reverse transcription (TPRT)", during which the endonuclease activity of ORF2 protein nicks the target DNA and the exposed 3'OH primes the polymerization reaction using L1 RNA as a template [322] (Fig. 1.9). Non-LRT retrotransposon sequences account for ~35% of the mammalian genome, including the long interspersed nuclear element (LINE) and short interspersed nuclear element (SINE). LINE-1 (L1) is the most common LINE and it alone comprises ~17% of the genomic DNA [324]. L1 is able to replicate autonomously and it also facilitates *in trans* the mobilization of SINEs and some cellular RNAs to produce processed pseudogenes (i.e. copies of genes that have lost their protein-coding ability). Here, I will review the general knowledge of A3 proteins as restriction factors against LTR and non-LTR retrotransposons.

1.5.1. A3-mediated restriction of LTR retrotransposons

Following the discovery of the anti-viral activity of A3 proteins against HIV, similar restrictive activities against murine ERVs, such as MusD and intracisternal A particle (IAP), have been reported [40, 165, 325, 326]. Similar to the paradigm of HIV restriction by A3s, where A3s need to get encapsidated into the budding virion through a RNA-dependent interaction with HIV Gag protein, direct interactions between A3 proteins and IAP Gag have been reported [326]. It is has been generally accepted that the inhibition of ERV requires the deaminase activity of A3s, as the characteristic G-to-A deamination footprints from A3G and A3F have been found in various recovered ERV genomes [328-330].
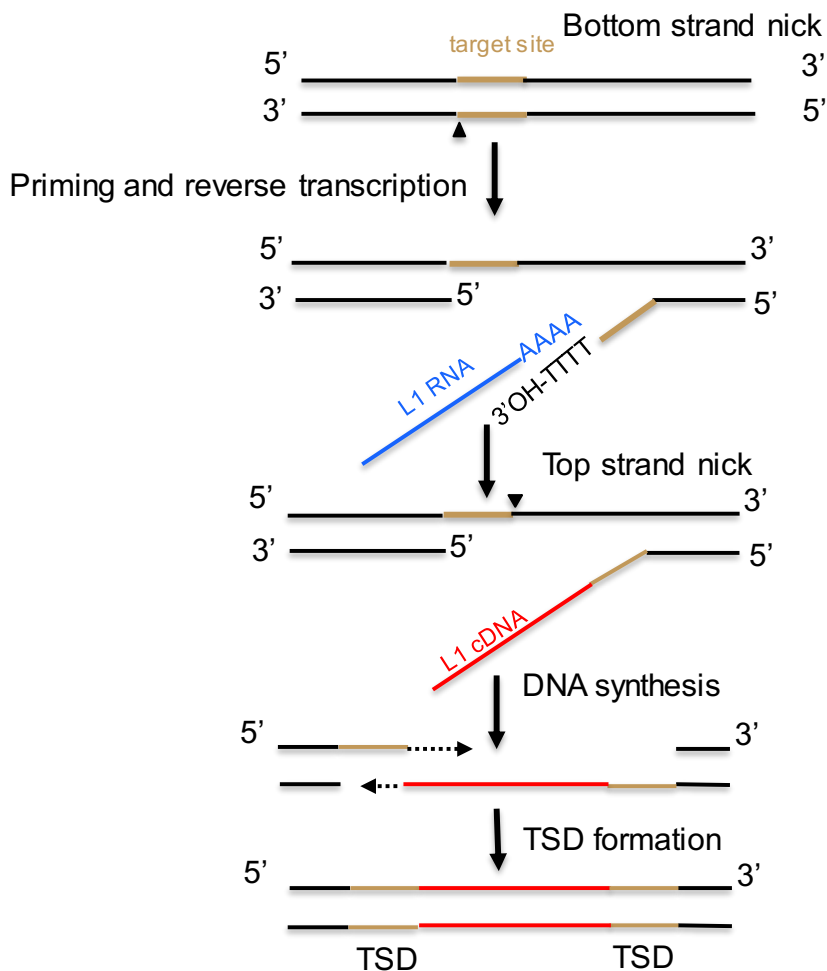
Figure 1.9 Schematic for target primed reverse transcription (TPRT) of L1. L1-encoded endonuclease generates a single-stranded nick on the genomic DNA at a degenerate concensus sequence (usually 5'-TTTT/A-3', with / represents the cleavage site). The exposed 3'OH is used by the L1 reverse transcriptase to prime L1 cDNA synthesis by paring the L1 RNA 3' polyA tail sequence from L1 RNA (blue) with a AT-rich sequence from the genomic DNA target site (wheat). L1 RNA is reverse transcribed by a L1-encoded reverse transcriptase to generate a cDNA copy (red). A second cleavage event is required to open the other strand of DNA, and microhomology between newly synthesized LINE-1 (-)cDNA and genomic DNA will prime (+)cDNA synthesis. TPRT ultimately leads to L1 integration at the new genomic location with the following structural hallmarks: target site duplications (TSDs), integration at an L1 endonuclease consensus cleave site and frequent 5' truncation in L1 sequences. This figure is re-drawn based on a figure in a review by Cordaux and Batzer [322].

C

G

↓ A3-catalyzed cytosine deamination

U

G

↓ Uracil DNA glycosylase

G

↓ AP endonuclease

3'OH   5'dRP

G

3'OH ↓ DNA polβ

3'OH

G

↓ DNA polβ and DNA ligase1

C

G
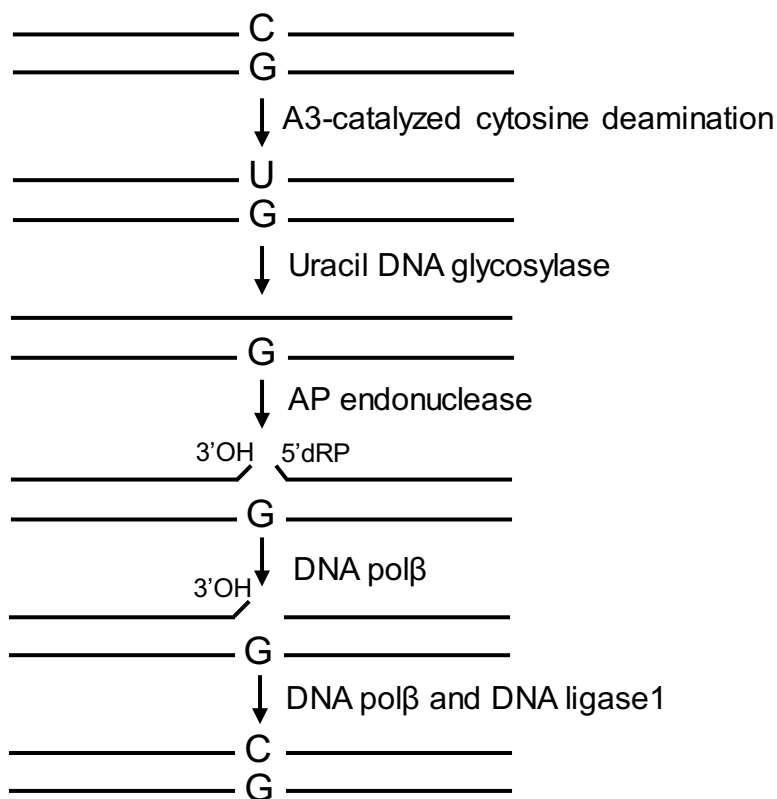
Figure 1.10 Repairing an uracil lesion in the DNA. The base excision repair (BER) pathway repairs small base distortions in the DNA helix structure, such as DNA damage from alkylation, oxidation and deamination. This schematic demonstrates how uracil lesion generated from A3's deamination activity is repaired. BER is initiated by uracil DNA glycosylase that recognizes and removes the uracil base. Uracil DNA glycosylase produces an abasic site on DNA by cleaving the bond between uracils and deoxyribose (i.e., the N-glycosidic bond). The abasic site is further processed by an apurinic/apyrimidinic (AP) endonuclease, a DNA polymerase and a ligase. The AP endonuclease cleaves the phosphodiester bond 5' to the abasic site, producing 3'OH and 5' deoxyribose phosphate (5'dRP) ends. The phosphodiesterase activity of Polβ, a major gap-filling family X polymerase, then cleaves the phosphodiester bond 3' to the lesion and removes 5'dRP and subsequently insert the correct nucleotide (a cytosine). The last step in BER is ligation of the nick after gap filling and DNA ligase1 is considered the major ligase to be involved in this process [327].

1.5.2. A3-mediated restriction of non-LTR retrotransposons

By overexpressing A3s in cell culture experiments, several studies have reported that different A3 family members are able to inhibit L1 to varying degrees, with A3A and A3B being the most potent at restriction [35, 42, 326, 331-334]. In general, a deamination-independent L1 restriction model has been proposed, as no G-to-A mutations related to A3-mediated deamination events have been detected. For a deamination-independent mechanism of L1 restriction, the A3s have been reported to interact with L1 mRNA and sequester L1 ribonucleoprotein (RNP) complex to a cytoplasmic compartment such as stress granules or physically interact with L1 RT and inhibit polymerization during L1 reverse transcription [42, 335, 336]. That the L1 restriction by A3 proteins does not seem to strictly correlate with their cellular localization to the nuclear compartment where L1 reverse transcription occurs supports the idea that A3s could inhibit various stages of the L1 life cycle [333]. It has been speculated that A3-mediated deaminations could still play a major role during L1 restriction, because the unsuccessful attempts to uncover A3-induced mutations were likely due to the degradation of uracil-containing cDNA intermediates by the uracil DNA glycosylase (UNG) and Apurinic/apyrimidinic (AP) endonuclease (APE)-mediated base excision repair machinery [103] [337] (for review, see [327]) (Fig. 1.10). Richardson *et al*. [337] proposed that the annealed L1 RNA normally protects the L1 (-)cDNA from deamination, but transiently exposed ssDNA would become accessible to A3A-mediated deamination during TPRT. The authors were able to recover G-to-A mutations in L1 replication intermediates in the presence of A3A when UDG was blocked [337]. Altogether, these data suggest the mechanisms of L1 restriction by A3s may involve both deaminase-dependent and independent activities.

2.0 Hypothesis and Objectives

2.1 Rationales and Hypothesis

Arthur Kornberg once said, "cell biologists look; geneticists count; biochemists clean". The A3 field is no exception to this, as the restrictive abilities of A3 proteins against HIV were originally discovered by a group of virologists and cell biologists across the world and it was not until careful inspection of the HIV genomic sequences did some interesting biochemical characteristics belonging to the cytidine deaminase activities of A3 enzymes become recognized. Building on the hypothesis that HIV Vif has developed a unique interaction with each A3 enzyme due to the different selection pressures A3s impose on HIV, we have characterized the biochemical activities of the single-domain A3H enzyme and compared that to the double-domain A3G, which has traditionally been the enzyme family representative. Based on this knowledge, we have also investigated the molecular mechanism of degradation-independent Vif-mediated inhibition of A3G and A3H deamination activity and the mechanisms that govern an efficient degradation-dependent Vif-mediated inhibition of these two enzymes. The central hypothesis of my Ph.D thesis is that knowledge gained on the inherent biochemical properties that enable A3G and A3H to locate and catalyze their deamination motifs would lead to a better understanding of the mechanistic basis of proviral hypermutations and the cellular activities of A3G and A3H as restriction factors against retroviruses (such as HIV) and retrotransposons (such as L1).

2.2 Objectives

The objectives of my Ph.D thesis include:

1. Investigation of the degradation-independent mode of inhibition imposed by Vif towards A3 enzymes through examination of the interaction between A3G and Vif and how Vif inhibits the deaminase activity of A3G.

2. Characterization and comparison of the interaction between A3G and A3H and different Vif variants and identification of common features that result in Vif-mediated degradation, which contribute to defining the determinants required for efficient Vif-mediated degradation of A3 enzymes.

3. Biochemical characterization and comparison of two stable A3H haplotypes and comparison of the oligomerization models for Z1, Z2 and Z3-type A3 enzymes.

4. Identification of the mechanisms A3H uses to restrict the replication of L1.

3.0 HIV-1 Viral infectivity factor alters processive single-stranded DNA scanning of the retroviral restriction factor A3G

Yuqing Feng[1], Robin P. Love[1] and Linda Chelico[1]
[1]Department of Microbiology & Immunology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada, S7N 5E5

Key words: Deoxycytidine deaminase, processivity, restriction factor, protein-protein interaction

Running title: Vif alters processive ssDNA scanning of APOBEC3G.

All experiments in this chapter were performed by Y.F.. Y.F. and L.C. conceived and designed the experiments. Y.F. and L.C., and R.P.L. analyzed the data. L.C. wrote the article. Y.F., R.P.L. and L.C. revised the article.

The information in this chapter was previously published:

3.1 Abstract

APOBEC3G is a retroviral restriction factor that can inhibit the replication of HIV-1 in the absence of the viral infectivity factor (Vif) protein. Virion encapsidated APOBEC3G can deaminate cytosine to uracil in viral (-) DNA which leads to hypermutation and inactivation of the provirus. APOBEC3G catalyzes these deaminations processively on single-stranded DNA using sliding and jumping movements. Vif is thought to primarily overcome APOBEC3G through an interaction that mediates APOBEC3G ubiquitination and results in its proteasomal degradation. However, Vif may also inhibit APOBEC3G mRNA translation, virion encapsidation, and deamination activity. Here we investigate the molecular mechanism of Vif_{IIIB}- and Vif_{HXB2}- mediated inhibition of APOBEC3G deamination activity. Biochemical assays using a model HIV-1 replication assay and synthetic single-stranded or partially double-stranded DNA substrates demonstrate that APOBEC3G has an altered processive mechanism in the presence of Vif. Specifically, Vif_{HXB2} inhibited the jumping and Vif_{IIIB} inhibited the sliding movements of APOBEC3G. The absence of such an effect by Vif on degradation resistant APOBEC3G D128K indicates that a Vif-APOBEC3G interaction mediates this effect. That the partially processive APOBEC3G was less effective at inducing mutagenesis in a model HIV-1 replication assay suggests that Vif co-encapsidation with APOBEC3G can promote sublethal mutagenesis of HIV-1 proviral DNA.

3.2 Introduction

APOBEC3G (A3G) is a single-stranded DNA (ssDNA) deoxycytidine deaminase that functions as a potent host restriction factor to inhibit HIV-1 replication in strains that lack the viral infectivity factor (Vif) [22-25]. In the absence of Vif, sufficient amounts of A3G are able to become incorporated into budding virions by binding to the viral genomic (+)RNA or 7SL cellular RNA, and/or nucleocapsid portion of the Gag [130]. Subsequently, A3G catalyzes the deamination of cytosine to uracil (C→U) on nascently formed viral (-)DNA in the target cell [90, 91]. HIV-1 reverse transcriptase (RT) uses these uracils as a template thereby creating G→A mutations on the viral genomic strand, resulting in hypermutation and inactivation of proviral DNA [22-24]. Synthesis of the (-)DNA is initiated from a host tRNA$^{\text{Lys,3}}$ primer that binds the HIV-1 primer binding site near the 5'-end of the viral genome [5]. Two polypurine tracts (PPT) that reside in the center and the 3'-end of the HIV-1 RNA genome are responsible for priming (+) DNA synthesis by virtue of their resistance to RNase H degradation [338]. The multiple priming

sites for (+) DNA synthesis allow for conversion of the single-stranded (-) DNA to a double-stranded (ds)DNA provirus as rapidly as possible and influences the distribution of A3G-induced mutations [339, 340].

A3G is able to search for its preferred 5'CC<u>C</u> or 5'C<u>C</u> deamination motif (underlined C is deaminated) using a processive scanning mechanism [132]. The scanning appears to be through facilitated diffusion [146, 148, 341] and involves sliding and microscopic jumping movements that enable a 3-dimensional search of the ssDNA without A3G diffusing into the bulk solution [132, 149, 150]. This searching mechanism gives enzymes the potential to increase their target searching efficiency [146, 148] and results in A3G-catalyzed deaminations occurring in a stochastic manner [342]. Further, since A3G must scan viral (-)DNA in the presence of short RNA/DNA hybrid regions that result from incomplete processing by the RT RNase H domain, the 3-dimensional search enables an efficient scan of the heterogenous substrate [139].

A3G contains two zinc-binding domains. The C-terminal domain (CTD) contains the catalytically active domain that is directly responsible for deaminase activity of A3G [56, 57] and can contribute to oligomerization [140, 343, 344]. The N-terminal domain (NTD) is catalytically inactive, but this half of the enzyme is responsible for processivity, virion encapsidation, and can also mediate oligomerization of A3G [56, 57, 137-139, 345]. The residues that appear to contribute to these N-terminal half functions are located on predicted loop 7, residues 124-130 [137-139] (Fig. 3.1).

A3G-mediated restriction of HIV-1 is for the most part inhibited by HIV-1 Vif [63, 70, 225, 246, 346]. Vif is thought to primarily accomplish this through triggering A3G degradation [63, 70, 225, 246]. Among the seven APOBEC3 family members (A3A, A3B, A3C, A3D, A3F, A3G and A3H), A3G appears to be the most potent at restricting HIV-1 [81, 171, 172], but also the most sensitive to Vif [295]. Specifically, Vif binds to A3G [63, 225, 246], is stabilized in host cells by binding the transcription cofactor core binding factor β (CBFβ) [68, 69], and recruits a cellular ubiquitin ligase complex [70]. This results in the polyubiquitination and degradation of A3G through the 26S proteasome pathway [247, 347]. Vif inhibits the antiviral activity of A3G in a species-specific manner [64]. For example, HIV-1 Vif fails to neutralize A3G from African green monkey or rhesus macaque and conversely simian immunodeficiency virus (SIV) Vif is incapable of neutralizing human A3G [64]. This species-specific type of A3G restriction has
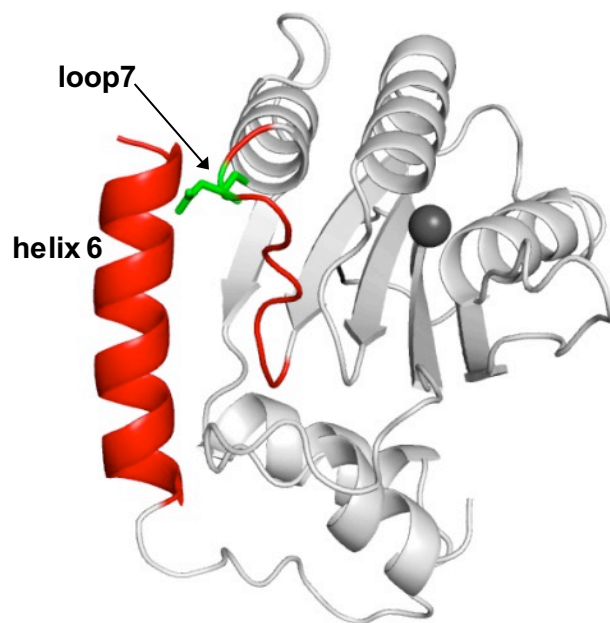
Figure 3.1 Model of the N-terminal half of A3G. Model (grey) shows loop 7 and helix 6 (both in red). The central residue that mediates the interaction with Vif, Asp128 is shown (green). Zinc atom is a dark grey sphere. The predicted model of A3G NTD was obtained by using the automated SWISS-MODEL program using the homologous A3G CTD (PDB: 3IQS) structure as a template. Figure was made using PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.5, Schrödinger, LLC.).

been correlated with the ability of Vif to physically associate with A3G and is centrally mediated by Asp[128] of A3G [280]. Mutating human A3G Asp[128] to Lys, as is found in African green monkey A3G, abrogates the effect of HIV-1 Vif on human A3G [278-280]. The Asp[128] is predicted to be located on loop 7 in the N-terminal half of A3G (Fig. 3.1) near the determinants for dimerization, processivity, and virion incorporation ([124]YYFW[127]) [137-139].

Additionally, Vif has been shown to inhibit A3G function in two other ways that appear to involve the RNA binding ability of Vif. First, Vif has been reported to promote A3G virion exclusion in a degradation-independent manner [346, 348, 349]. A3G needs to bind to the viral genomic RNA or cellular 7SL RNA, and/or nucleocapsid portion of HIV-1 Gag in order to be encapsidated in the budding virus [130]. It has been suggested that Vif recruits A3G into high molecular weight RNA/protein masses that are unable to be encapsidated, but the exact mechanism has not been identified [296]. Vif may exert this function as a secondary effect of its own attempt to be encapsidated into budding virions, although the amount of Vif estimated in virions ranges widely from about 7 to 80 molecules [94, 95, 223, 350]. Second, there is evidence to suggest that Vif depletes the intracellular pool of A3G by downregulating its mRNA translation and stability [223, 225, 293, 294]. It is thought that Vif can directly bind A3G mRNA to exert this function [293, 294].

Despite the multiple ways Vif attempts to inhibit A3G virion encapsidation, Vif neutralization of A3G may not be absolute. A3G may be capable of escaping Vif inhibition in the virus producing cell, resulting in A3G encapsidation into budding virions [124, 223]. Nowarski *et al.* reported that a range of 0.3-0.8 molecules of A3G can be found in wild type (wt) HIV-1 virions, far less than the 4-9 molecules of A3G found in Vif-deficient virions [93]. Further, the residual A3G encapsidated with Vif in the virion was found to be less catalytically active than A3G encapsidated in its absence [298]. The explanation of this observation is that virus encapsidated Vif inhibits A3G deamination activity [298]. This is supported by an earlier report which characterized this phenomenon when A3G and Vif were co-expressed in *E. coli* [299]. The mechanism by which Vif inhibits A3G deamination activity is not fully understood. Yet, this mode of Vif-mediated inhibition has implications for design of Vif-based HIV-1 therapeutics that exclude Vif from interacting with ubiquitin ligase components [300, 311] but allow it to remain in contact with A3G and the relationship between A3G deamination activity

and sublethal mutagenesis of HIV-1 [73, 124, 126, 351]. Vif-mediated partial inhibition of A3G deamination activity could potentially lead to viral evolution rather than viral inactivation.

Here we investigate the molecular mechanism of degradation-independent Vif-mediated inhibition of A3G deamination activity. Two Vif variants, Vif_{HXB2} and Vif_{IIIB} have been applied to our study. We used Vif_{IIIB} as a standard variant from a widely used HIV-1 lab isolate and we used Vif_{HXB2} since it has been reported to be more potent at Vif-mediated degradation of APOBEC3 enzymes [295]. The data indicate that the Vif variants inhibit A3G deamination activity by primarily altering its processive scanning mechanism, which decreases the ability of A3G to induce mutations in nascently reverse transcribed cDNA. Each Vif variant can inhibit A3G processive scanning in a unique way suggesting that there are variant specific interactions that underlie the differential potency of Vif_{HXB2} and Vif_{IIIB} against APOBEC3 enzymes.

3.3 Results

3.3.1 Decreased A3G-induced mutagenesis in the presence of Vif. A reconstituted model HIV-1 replication system was used to determine whether Vif could inhibit the deamination ability of A3G. In this assay, we used an RNA construct that contains (from the 5'- to 3'-end) a PPT, 120 nt of the HIV-1 *prot* active site, and the *lacZα* gene. The PPT enables second strand synthesis to occur. The *prot* and *lacZα* sequences are used to characterize the spectra of mutations. The *prot* sequence is further used to gauge the ability of A3G-induced deaminations to inactivate the protease enzyme. As such, the sequence is analyzed in the (+) strand orientation where C→U deaminations are detected as G→A mutations. A3G is able to cause approximately a 17-fold increase in the population mutation frequency and 11-fold increase in the clone mutation frequency as compared to RT alone (compare Table 3.1 and Table 3.2).

A plot of the distribution of A3G-induced mutations recovered from the sequenced clones demonstrates a mutational gradient in the 5'→3' orientation, meaning that deaminations were biased towards the 5'-end of the nascently synthesized cDNA (Fig. 3.2A). This has been previously observed for A3G *in vivo* due to HIV-1 replication dynamics in which (-)DNA furthest from the PPT is single-stranded the longest and incurs the most deaminations [91, 339, 340]. To determine if Vif could influence the deamination activity of A3G we added Vif at excess (~2-fold more than A3G). It can be seen by visual inspection of the mutation spectra that addition of Vif_{IIIB} or Vif_{HXB2} causes a decrease in A3G-induced mutagenesis (Fig 3.2A), with Vif_{HXB2} (Fig. 3.2B) having a greater inhibitory effect than Vif_{IIIB} (Fig. 3.2C). The decrease in the

Table 3.1 A3G-mediated mutation frequencies in a model HIV replication system in the absence and presence of Vif$_{HXB2}$ and Vif$_{IIIB}$.

Reaction condition	Population mutation frequency	Clone mutation frequency ($\times10^{-2}$)
A3G	0.87	2.2
A3G + Vif$_{HXB2}$	0.63[a]	0.9[a]
A3G + Vif$_{IIIB}$	0.90	1.5[b]

The ratio of white colonies to total colonies is defined as the population mutation frequency. The average number of G→A mutations per base pair in the 368 nt *prot-lacZα* construct is defined as the clone mutation frequency.
[a] Significant difference was designed as p≤0.001 versus A3G values.
[b] Significant difference was designed as p≤0.01 versus A3G values.
[c] Significant difference was designed as p≤0.05 versus A3G values.

Table 3.2 Mutation frequencies in a model HIV replication system in the absence of A3G

Reaction condition	Population mutation frequency	Clone mutation frequency (x10^{-2})
No A3G	0.05	0.2

The ratio of white colonies to total colonies is defined as the population mutation frequency. The average number of mutations (substitutions, insertions or deletions) per base pair in the 368 nt *prot-lacZα* construct is defined as the clone mutation frequency.
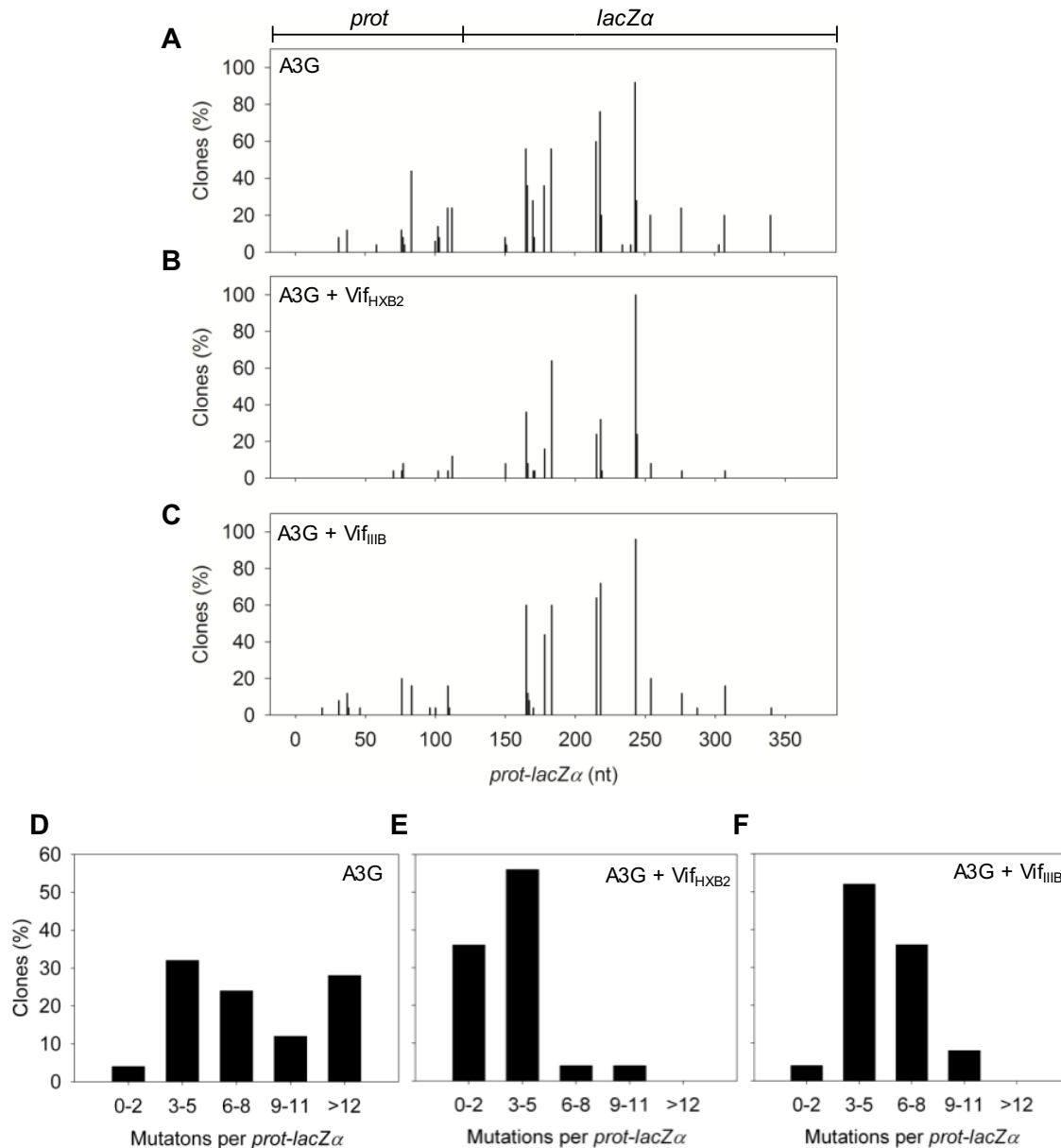
Figure 3.2 A3G-induced mutagenesis is inhibited by Vif$_{HXB2}$ and Vif$_{IIIB}$ in a model HIV replication system. A-C, spectra of mutations incurred in 368 nt prot-lacZα construct are plotted as percent of clones containing a mutation at a particular location (nt) for (A) A3G, (B) A3G in the presence of Vif$_{HXB2}$ and (C) A3G in the presence of Vif$_{IIIB}$. D-F, histograms were generated to illustrate the population distribution of mutations per prot-lacZα construct for (D) A3G, (E) A3G in the presence of Vif$_{HXB2}$ and (F) A3G in the presence of Vif$_{IIIB}$.

A3G-induced mutations when Vif_{HXB2} (Fig. 3.2B) or Vif_{IIIB} (Fig. 3.2C) is present is greater in the *prot*, likely owing to the shorter time that this region is single stranded in comparison to the *lacZα* due to the proximity to the PPT.

The clones were also binned for analysis based on the number of mutations per *prot-lacZα* (Fig. 3.2, D-F). A3G is able to induce a widespread number of mutations (Fig. 3.2D). We found a range of 3 to 12 mutations per *prot-lacZα* occurring for A3G. The data are far less spread when A3G-induced mutagenesis occurs in the presence of Vif_{HXB2} where we find there are only 3-5 mutations per *prot-lacZα* ~55% of the time and 0-2 mutations per *prot-lacZα* ~35% of the time (Fig. 3.2E). For Vif_{IIIB} the reduction in A3G-induced mutations is demonstrated by a shift of mutations down to 3 to 8 mutations per *prot-lacZα* (Fig. 3.2F). The data were analyzed further by comparing the frequency of mutations in heavily mutated sites in the presence of A3G and the absence and presence of Vif (Table 3.3). From this analysis we found that Vif_{HXB2} causes 6 of the 8 heavily mutated sites to be mutated at a lower level (Table 3.3, A3G + Vif_{HXB2}). The detrimental effect of Vif_{IIIB} on A3G-induced mutagenesis was only significant in the *prot* region (Table 3.3, A3G + Vif_{IIIB}).

The more potent effect of Vif_{HXB2} is further evident when the mutation frequencies are examined. Here we scored the number of white colonies in the population (population mutation frequency) or the number of mutations per base pair (clone mutation frequency) (Table 3.1). We found that Vif_{HXB2} causes a significant decrease in both the population (1.4-fold) and clone mutation frequency (2-fold) from A3G alone (Table 3.1). The influence of Vif_{IIIB} was not as large with the population mutation frequency remaining the same as for A3G alone and the clone mutation frequency decreasing 1.5-fold (Table 3.1). Altogether, the data indicate that both Vif variants we tested can inhibit A3G-induced mutagenesis, consistent with previous results which tested Vif_{IIIB} [299] or Vif_{HXB2} [298]. Further, the effect of Vif_{HXB2} on A3G is more potent than Vif_{IIIB}, which is consistent with how these Vif variants function with respect to inducing degradation of A3G and other APOBEC3 enzymes [295].

3.3.2 Vif-mediated inhibition of A3G can promote sublethal mutagenesis in a model HIV-1 replication system. Here we assessed the potential ability of each Vif variant to cause A3G to induce a sublethal amount of mutagenesis in the *prot* (Table 3.4). After determining the protease amino acid sequence of the mutated clones we can infer whether the protease would retain activity or be inactivated by drawing on data from Loeb *et al.* [352]. in which an extensive

Table 3.3 Frequency of A3G hot-spot mutations in the *prot-lacZα*.

Site (nt)	Frequency of mutation		
	A3G	A3G + Vif$_{HXB2}$	A3G + Vif$_{IIIB}$
85	0.44	0.00^a	0.16^b
111	0.24	0.04^b	0.16^c
114	0.24	0.12^c	0.00^b
166	0.56	0.36^b	0.60
184	0.56	0.64^c	0.60
216	0.60	0.24^b	0.64
219	0.76	0.32^a	0.72
244	0.92	1.00	0.96

A hot-spot was defined as occurring in at least 20% of clones for the *prot* region (1 - 120 nt) and at least 50% of clones for the *lacZα* region (121 - 368 nt).
[a] Significant difference was designed as $p \leq 0.001$ versus A3G values.
[b] Significant difference was designed as $p \leq 0.01$ versus A3G values.
[c] Significant difference was designed as $p \leq 0.05$ versus A3G values.

Table 3.4 Amino acid changes to the HIV-1 protease resulting from A3G-mediated deaminations in the absence and presence of Vif$_{HXB2}$ and Vif$_{IIIB}$.

Amino acid	Codon change	Amino acid change	Predicted protease activity	Mutated clones (%)		
				A3G	A3G + Vif$_{HXB2}$	A3G + Vif$_{IIIB}$
21	GAA → AAA	E → R	+	0	0	4
25	GAT → AAT	D → N	+	8	0	8
27	GGA → AGA	G → R	-	12	0	12
30	GAT → AAT	D → N	+	0	0	4
34	GAA → AAA	E → K	+	4	0	0
38	TTG → TTA	L → L	+	0	4	0
40	GGG → AGG	G → R	-	12	4	20
	GGG → GAG	G → E	-	8	8	0
	GGG → GGA	G → G	+	0	4	0
42	TGG → TAG	W → stop	-	44	0	16
46	ATG → ATA	M → I	+	0	0	4
48	GGG → AGG	G → R	+	4	0	0
	GGG → AGA	G → R	+	4	0	0
	GGG → GGA	G → G	+	12	0	4
49	GGA → AGA	G → R	-	4	0	0
51	GGA → AGA	G → R	-	24	4	16
	GGA → GAA	G → E	-	4	0	4
52	GGT → AGT	G → S	-	24	12	0

Predicted protease activity is determined from a mutagenesis study conducted by Loeb *et al[352]*. An active protease is denoted with a + and an inactive protease is denoted with a -.
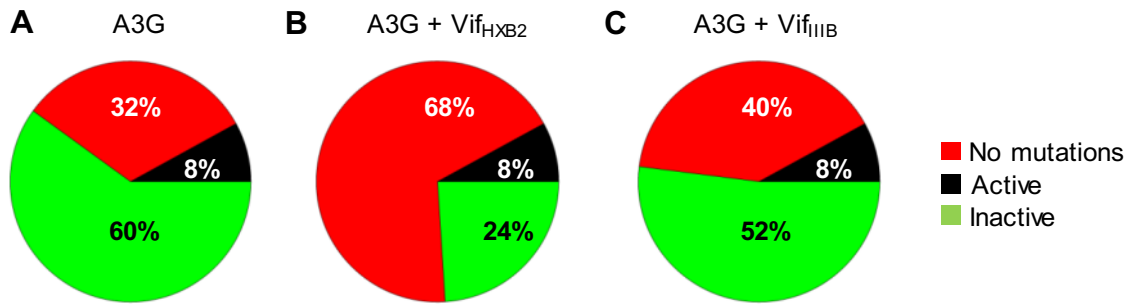
Figure 3.3 Vif_{HXB2} and Vif_{IIIB} have a different effect on the decrease in A3G-induced mutations in the prot sequence. Individual analysis of each prot clone enabled determination of the percent of clones that would result in an unmutated (no mutations, red), mutated and active (black), or mutated and inactive (green) prot. A, A3G was able to inactivate the prot in 60% of clones and left an active prot in 8% of mutated clones. A3G did not induce any mutations in the prot in 32% of clones. B, A3G-induced mutagenesis in the presence of Vif_{HXB2} resulted less mutations overall. There were 68% of clones with no mutations, 24% of clones were mutated and inactive, and 8% of clones were mutated but retained prot activity. C, the effect of Vif_{IIIB} on A3G-induced mutagenesis was less than Vif_{HXB2} and resulted in 40% of clones with no mutations, 52% of clones with inactive prot, and 8% of clones with mutated but active prot.

mutagenesis of the protease was conducted. Concomitant with the site to site decreases in mutation frequency (Table 3.3), we saw a decrease in the number of clones with amino acid changes of the protease in the presence of the Vif variants (Table 3.4).

When A3G-catalyzed deaminations occurred in the presence of Vif_{HXB2}, the types of mutations incurred were not changed, but rather decreased with most mutations seen with A3G alone being absent in the presence of Vif_{HXB2} (Table 3.4). In the presence of Vif_{IIIB} there were decreases in mutations at 9 of the 13 sites mutated for A3G alone and changes in the mutation frequency, i.e., increased mutation or novel site in 4 of the 13 sites mutated for A3G alone (Table 3.4). For example, the most heavily mutated site for native A3G is at amino acid 42 (Trp to stop; 44%). However, in the presence of Vif_{HXB2} and Vif_{IIIB}, the frequency of this mutation drops to 0 % and 16%, respectively (Table 3.4). It is interesting that some of the sites that were mutated only in the presence of Vif_{IIIB}, albeit at a low level, cause protease inhibitor resistance (Table 3.4, 4% for D30N and M46I).

To determine if each individual model provirus synthesized in our assay would code for an active or inactive protease we analyzed each clone individually (Fig. 3.3). A3G-induced mutagenesis is able to cause inactivation of the *prot* 60% of the time and leaves the *prot* active only 8% of the time (Fig. 3.3A). A significant amount of clones (32%) had no mutations in the *prot* region. A3G–induced mutagenesis in the presence of Vif_{HXB2} changes this distribution so that 68% of clones had no mutations in the *prot* and 24% had mutations that inactivated the *prot* (Fig. 3.3B). Again, 8% of clones were mutated but still coded for an active *prot* (Fig. 3.3B). The same trend was found for A3G in the presence of Vif_{IIIB} where slightly more clones were not mutated (Fig. 3.3C, 40%) in comparison to A3G alone (Fig. 3.3A, 32%). These results are in agreement with the analysis of the total population of mutations (Table 3.4) where there was an overall decrease in mutations, rather than an alteration of the mutational hot spots.

3.3.3 Vif alters the processive scanning mechanism of A3G. To investigate how Vif_{IIIB} and Vif_{HXB2} inhibit A3G-induced mutagenesis and why these variants differ in their effect (Fig. 3.3, Table 3.1 and 3.3), we conducted assays measuring A3G-catalyzed deaminations on synthetic DNA substrates. Specifically, the assay can measure the processive scanning behavior of A3G by using a 118 nt ssDNA substrate with two 5'CCC deamination motifs spaced 61 nt apart (Fig. 3.4A, *schematic*). Figure 3.3A shows a characteristic result of A3G in this assay. A3G deaminates both the 5'- and 3'- proximal CCC motifs, but there is a 2-fold preference for the 5'-

66

A

5' ——15 nt—— CCC ——61 nt (F)—— CCC ——36 nt—— 3'

DNA only / A3G / + Vif$_{HXB2}$ / + Vif$_{IIIB}$

5' C
3' C
5' C & 3' C

←118 nt
←100 nt
←81 nt
←63 nt

Processivity factor   9.1 4.1 5.6

B

33 nt   8 nt

5' ——15 nt—— CCC ——20 nt DNA (F)—— CCC ——36 nt—— 3'

DNA only / A3G / + Vif$_{HXB2}$ / + Vif$_{IIIB}$

5' C
3' C
5' C & 3' C

←118 nt
←100 nt
←81 nt
←63 nt

Processivity factor   4.5 1.1 5.5

C

33 nt   8 nt

5' ——15 nt—— CCC ——20 nt RNA (F)—— CCC ——36 nt—— 3'

DNA only / A3G / + Vif$_{HXB2}$ / + Vif$_{IIIB}$

5' C
3' C
5' C & 3' C

←118 nt
←100 nt
←81nt
←63 nt

Processivity factor   4.8 1.0 4.1

67

Figure 3.4 Processivity of A3G in the absence and presence of Vif$_{HXB2}$ and Vif$_{IIIB}$. A, deamination of a 118 nt F-labeled ssDNA substrate by A3G. Two CCC motifs within the ssDNA sequence are spaced 61 nt apart. Single deaminations of the 5'C and 3'C are detected as the appearance of labeled 100- and 81- nt fragments, respectively; double deamination of both C residues on the same molecule results in a 63 nt labeled fragment (5'C & 3'C). A3G processivity is decreased by ~2-fold in the presence of Vif$_{HXB2}$ (+Vif$_{HXB2}$) and Vif$_{IIIB}$ (+Vif$_{IIIB}$). B, deamination of the substrate described for (A), but with a 20 nt ssDNA annealed between the two CCC motifs which can block the sliding component of A3G processivity. This results in a 2-fold decrease in A3G processivity. Processive deaminations are still observed due to the jumping component of A3G processive movement. In the presence of Vif$_{HXB2}$ the processivity of A3G is nearly absent indicating that the jumping component of A3G has been inhibited. This is in contrast to the Vif$_{IIIB}$ that has no effect on A3G jumping movement. C, deamination of the substrate described for (B) but with a 20 nt RNA annealed between the two CCC motifs. Processive scanning of A3G on this substrate in the absence and presence of Vif$_{HXB2}$ and Vif$_{IIIB}$ was similar to that observed for (B). The measurements of processivity (Processivity factor) are shown below the gel. Values are an average from at least three independent experiments and the SEM for the processivity factors are (A) A3G, 1.90; +Vif$_{HXB2}$, 0.03; +Vif$_{IIIB}$, 0.65 (B) A3G, 0.80; +Vif$_{HXB2}$, 0.20; +Vif$_{IIIB}$, 0.76; (C) A3G, 1.20; +Vif$_{HXB2}$, 0.16; +Vif$_{IIIB}$, 0.40.

motif due to an inherent catalytic orientation specificity of the active site [141]. Of importance is the presence of a double deamination band which indicates that a deamination occurred at both the 5'C and 3'C on the same ssDNA (Fig. 3.4A, 63 nt). The processivity factor is calculated as a ratio of the intensity of this band to the calculated expected value of deaminations that would occur at both 5'- and 3'- proximal motifs if the enzyme were not processive (*see Material and Methods*). The processivity factor of 9.1 (Fig. 3.4A, below gel) means that A3G is 9 times more likely to undergo a processive deamination of both CCC motifs than deaminate each of the motifs in separate enzyme-ssDNA encounters. Vif_{HXB2} and Vif_{IIIB} are able to decrease the processivity of A3G by at least 1.5-fold from 9.1 to 4.1 or 5.6, respectively (Fig. 3.4A).

To discern how the Vif variants are able to decrease the processivity of A3G we used an assay that can make a distinction between the two scanning modes of A3G. A3G has been shown to scan ssDNA by sliding and microscopic dissociations and reassociations termed jumping [132, 139]. This bifunctional processive mode was identified using an experiment where a complementary DNA oligonucleotide is annealed between two CCC motifs on an ssDNA substrate (Fig. 3.4B, *schematic*). The partially dsDNA, to which the enzyme cannot effectively bind, acts as a block to the sliding component of A3G processivity [132, 141]. However, A3G can still transverse the block by jumping. As such, A3G incurs a ~2-fold decrease in processivity (compare Fig. 3.4*A* and 3.4B, processivity factors of 9.1 and 4.5, respectively). When Vif_{HXB2} is added to the reaction with the partially dsDNA substrate, the processivity factor of A3G decreases another 2-fold, from 4.1 to 1.1 (Fig. 3.4, A and B, processivity factor) and there is a visible decrease of the double deaminations to almost none (Fig. 3.4B, 63 nt). This indicates that Vif_{HXB2} can inhibit the jumping component of A3G so that it is no longer able to transverse the dsDNA. A3G is still able to slide since processive deaminations occurred on fully ssDNA (Fig. 3.3A). Vif_{IIIB} does not affect the ability of A3G to jump since the processivity factor in the presence of the dsDNA is 5.5 (Fig. 3.4B, processivity factor), which is similar to the processivity of A3G in the presence of Vif_{IIIB} on fully ssDNA (Fig. 3.4A, processivity factor of 5.6). However, Vif_{IIIB} does appear to affect A3G processivity since it decreased the processivity factor on ssDNA 1.5-fold (Fig. 3.4A, processivity factor decreased from 9.1 to 5.6), suggesting that Vif_{IIIB} may inhibit A3G sliding. Similar results were found when we used a complimentary RNA oligonucleotide as a block (Fig. 3.4*C*). An RNA/DNA hybrid block would be what A3G

Figure 3.5 Mutation cluster frequency of A3G is decreased in the presence of Vif$_{IIIB}$ but not Vif$_{HXB2}$. Sliding movement of A3G is inferred from clustered mutations. Jumping movement of A3G is inferred from distantly spaced mutations. Frequency of clustered mutations was scored in eight regions of the prot-lacZα that contain three to six G residues. Horizontal bars represent the average mutation cluster frequency. The data show that A3G alone and in the presence of Vif$_{HXB2}$ induce a similar frequency of clustered mutations (p value 0.32). However, A3G in the presence of Vif$_{IIIB}$ induced significantly less clustered mutations (p value 0.02).

encounters on HIV-1 (-)DNA during viral replication and suggests that the decrease in A3G-induced mutagenesis seen in the HIV-1 replication assay (Fig. 3.2) was due to a partial inhibition of A3G processivity by the Vif variants (Fig. 3.4, *A-C*).

To further test this hypothesis, we examined the sequences of clones from the HIV-1 replication assay. We can analyze the spatial proximity of mutations in the clones and relate it to processivity where clustered mutations are indicative of scanning by local sliding and distantly spaced singleton mutations are indicative of scanning by jumping [139]. Areas of the 368 nt (+)DNA sequence (contains *prot* and *lacZα*), which contain multiple G residues were used to analyze the frequency of clustered A3G-induced G→A mutations. If the sliding component of A3G ssDNA scanning mechanism is retained, A3G should be able to induce mutagenesis at multiple G residues in a close region, i.e., 10 nt or less. The mean clustered mutation frequency for A3G alone (Fig. 3.5, 0.48) is similar to the situation when Vif_{HXB2} is present (Fig. 3.5, 0.33, p value 0.32). These data are consistent with Fig. 3.4 *A-C*, which shows that A3G retains the ability to slide, but not jump, in the presence of Vif_{HXB2}. In the presence of Vif_{IIIB}, A3G is less capable of inducing mutations that are closely spaced (Fig. 3.5, 0.14, p value 0.02). This indicates A3G is less able to scan ssDNA by sliding when Vif_{IIIB} is present and is in agreement with processivity data obtained from synthetic substrates (Fig. 3.4, *A-C*).

3.3.4 Inhibitory mechanism of Vif is mediated by an interaction with A3G. A3G is known to physically interact with Vif [63, 225, 246], which suggests that Vif could alter the processivity of A3G through a protein-protein interaction. That the two Vif variants affect A3G processivity differently supports this notion (Fig. 3.4 and Fig. 3.5). To investigate this possibility, we used the A3G D128K mutant (referred to as D128K) which is insensitive to Vif-mediated degradation and has been shown to have a disrupted physical interaction with Vif [353]. Our rotational anisotropy data confirm the previous findings as we found no detectable interaction between fluorescein (F) labeled D128K and the Vif variants (Table 3.5). However, using F-A3G we found that Vif_{HXB2} and Vif_{IIIB} interact similarly with F-A3G. The measured apparent K_d values of 90 nM (Vif_{HXB2}) and 78 nM (Vif_{IIIB}) for F-A3G indicate a strong interaction can occur between these proteins (Table 3.5).

We tested the ability of D128K to induce mutations in a model HIV-1 replication system in the absence and presence of the Vif variants. The mutational spectrum for D128K alone shows

Table 3.5 Apparent dissociation constants (K_d) of A3G, D128K, and Vif variants.

Enzyme	Apparent dissociation constant (K_d), nM			
	F-A3G	F-D128K	*prot* (-) ssDNA	118 nt ssDNA
A3G			270 ± 59	203 ± 18
D128K				183 ± 13
Vif$_{HXB2}$	90 ± 15	no binding detected	27 ± 1	67 ± 18
Vif$_{IIIB}$	78 ± 10	no binding detected	92 ± 13	79 ± 15
A3G + Vif$_{HXB2}$				158 ± 20
A3G + Vif$_{IIIB}$				135 ± 11
D128K + Vif$_{HXB2}$				72 ± 10
D128K + Vif$_{IIIB}$				51 ± 19

The apparent K_d values are shown with the standard error of the mean that was calculated from three independent experiments. No value indicates that the experiment was not conducted.

Table 3.6 D128K-mediated mutation frequencies in a model HIV replication system in the absence and presence of Vif_{HXB2} and Vif_{IIIB}.

Reaction condition	Population mutation frequency	Clone mutation frequency ($x10^{-2}$)
D128K	0.96	2.2
D128K + Vif_{HXB2}	0.93	2.3
D128K + Vif_{IIIB}	0.90	2.6[c]

The ratio of white colonies to total colonies is defined as the population mutation frequency. The average number of G→A mutations per base pair in the 368 nt *prot-lacZα* construct is defined as the clone mutation frequency
[a] Significant difference was designed as $p \leq 0.001$ versus D128K values.
[b] Significant difference was designed as $p \leq 0.01$ versus D128K values.
[c] Significant difference was designed as $p \leq 0.05$ versus D128K values.
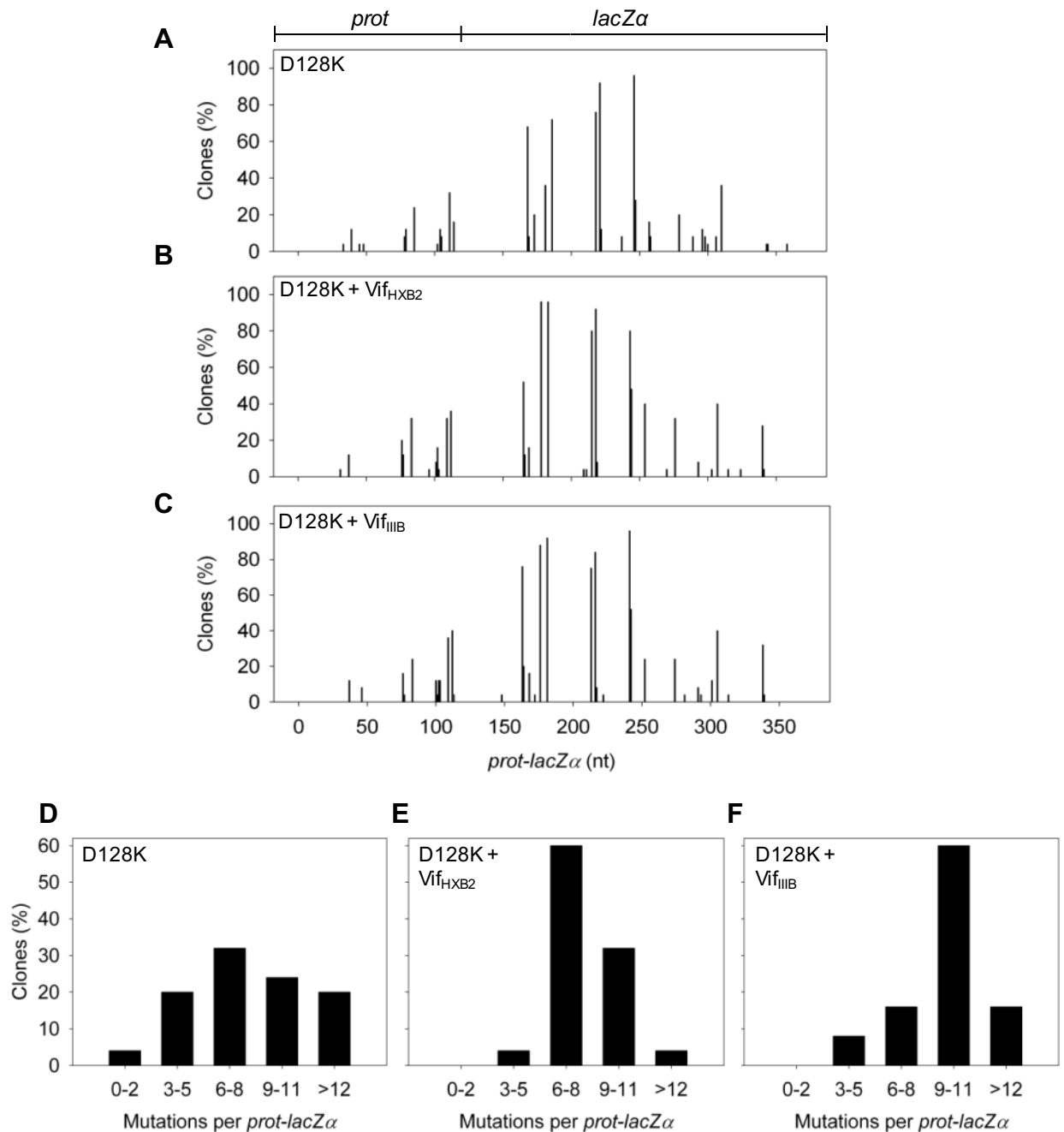
Figure 3.6 A3G D128K-induced mutagenesis is unaffected by Vif_{HXB2} and Vif_{IIIB} in a model HIV replication system. A-C, spectra of mutations incurred in 368 nt prot-lacZα construct are plotted as percent of clones containing a mutation at a particular location (nt) for (A) D128K, (B) D128K in the presence of Vif_{HXB2} and (C) D128K in the presence of Vif_{IIIB}. D-F, histograms illustrate that the population distribution of mutations per prot-lacZα construct for (D) D128K which is not decreased in the presence of (E) Vif_{HXB2} or (F) Vif_{IIIB}.

peaks for highly mutated sites in the same locations as for A3G (compare Fig. 3.2*A* and Fig. 3.6*A*). The intensity of the mutations among the sites differs slightly, likely owing to the stochastic nature of A3G deamination. The mutation frequencies are nearly identical for D128K and A3G at the population level (Table 3.6, D128K, 0.96; Table 3.1, A3G, 0.87) and clone level (Table 3.6, D128K, 2.2×10^{-2}; Table 3.1, A3G, 2.2×10^{-2}). Overall, the results are consistent with D128K behaving similar to A3G in this assay. When both Vif_{IIIB} and Vif_{HXB2} are added to the reaction with D128K we observe no decrease in the mutagenic ability of D128K (Fig. 3.6, *A-F* and Table 3.6), in contrast to A3G (Fig. 3.2, *A-F* and Table 3.1). In addition, we found no effects on D128K processive scanning by sliding by the addition of Vif_{HXB2} or Vif_{IIIB} to the model HIV-1 replication assay as the mutation cluster frequency (Fig. 3.7) and frequency of mutations (Fig 3.6 and Table 3.6) remained constant in all three experimental conditions. This observation is in agreement with data obtained on the 118 nt substrate where neither Vif_{HXB2} or Vif_{IIIB} were found to change D128K processivity (data not shown). Altogether, the data support a model in which an alteration in the processivity of A3G in the presence of Vif is caused by a Vif-A3G interaction. *3.3.5 Vif affects the specific activity of A3G and D128K*. Despite evidence that a Vif-A3G interaction is responsible for the observed decrease in A3G-induced mutagenesis (Table 3.4 and compare Fig. 3.5 and Fig. 3.2), Vif has been shown to bind DNA and RNA molecules [293, 354, 355]. Therefore, we assessed whether the Vif variants could additionally alter A3G deamination activity, i.e., decrease specific activity, by binding to the ssDNA substrate. Rotational anisotropy data show that A3G could bind to the substrate used in the processivity assays (Fig. 3.4*A*, 118 nt substrate) with an apparent K_d of 203 nM (Table 3.5), which is approximately 3-fold higher than the apparent K_d values of the Vif variants for this substrate (Table 3.5, K_d of Vif_{HXB2}, 67 nM; K_d of Vif_{IIIB}, 79nM). We also examined the binding of A3G and Vif to the *prot* (-)DNA which was part of the deamination substrate for the HIV-1 replication assay (Fig. 3.2). We observed A3G binding to the *prot* (-)DNA (Table 3.5, K_d of A3G, 270 nM) with an apparent K_d that was 10-fold (Table 3.5, K_d of Vif_{HXB2}, 27 nM) or 3-fold (Table 3.5, K_d of Vif_{IIIB}, 92 nM) higher than the Vif variants binding to this substrate. Therefore, in both our experimental systems, the apparent K_d values of Vif for DNA and F-A3G (Table 3.5, K_d of F-A3G from Vif_{HXB2} and Vif_{IIIB} is 90 nM and 78 nM, respectively) are similar, giving the possibility that Vif could compete with A3G for the ssDNA substrate and decrease the specific activity of A3G.

To test this hypothesis, we used D128K, which does not interact with Vif (Table 3.5).
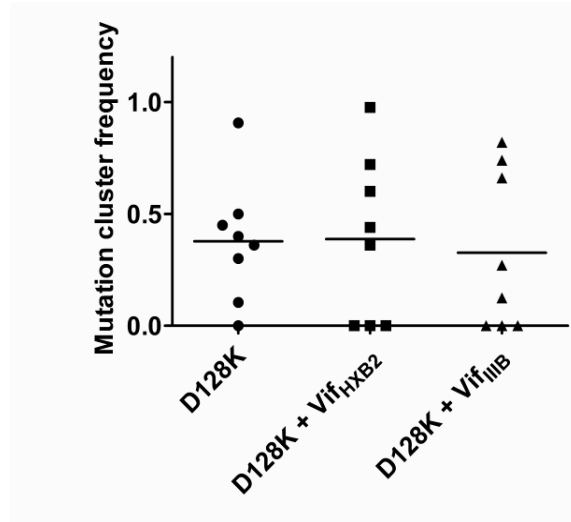
Figure 3.7 Vif$_{IIIB}$ and Vif$_{HXB2}$ do not decrease the mutation cluster frequency of D128K. Analysis of sequenced clones obtained from HIV replication assay conducted for D128K in the absence and presence of Vif$_{HXB2}$ and Vif$_{IIIB}$. Frequency of clustered mutations was scored in eight regions of the prot-lacZα that contain three to six G residues. Inferred sliding and jumping movement from clustered and distantly spaced mutations, respectively, are unchanged by the presence of Vif$_{HXB2}$ or Vif$_{IIIB}$. Horizontal bars represent the average mutation cluster frequency. The data show that D128K alone and in the presence of Vif$_{HXB2}$ or Vif$_{IIIB}$ induce a similar frequency of clustered mutations (p value 0.96 and 0.75, respectively).

Table 3.7 Specific activities of A3G and D128K in the absence and presence of Vif variants.

Reaction condition	Specific Activity (pmol/µg/min)
A3G	10 ± 0.3
A3G + Vif_{HXB2}	1.2 ± 0.3
A3G + Vif_{IIIB}	1.8 ± 0.9
D128K	40 ± 9
D128K + Vif_{HXB2}	0.5 ± 0.2
D128K + Vif_{IIIB}	0.8 ± 0.4

The specific activity was determined using the 118 nt substrate and the values are shown with the standard error of the mean that was calculated from three independent experiments.

There was no effect of Vif on D128K-induced mutagenesis (Fig. 3.7), suggesting that Vif interactions with the DNA in the HIV-1 replication assay are inconsequential with regards to processivity (Fig. 3.6). Notwithstanding, a decrease in specific activity could be occurring, but cannot be observed in the HIV-1 replication assay because deamination events are dependent on and limited by other factors such as the rates of RNA degradation and (+)DNA synthesis. To determine if Vif can decrease the specific activity of A3G or D128K, we used the 118 nt ssDNA substrate previously used for the processive deamination assay (Fig. 3.4). The experiment involved Vif and D128K or A3G being preincubated together and then the reaction was started by the addition of the 118 nt ssDNA substrate. On the 118 nt ssDNA substrate, D128K has a 4-fold higher specific activity (Table 3.7, 40 pmol/μg/min) than A3G (Table 3.7, 10 pmol/μg/min). However, Vif (HXB2 and IIIB) can decrease D128K specific activity 80-fold (Table 3.7, D128K + Vif$_{HXB2}$, 0.5 pmol/μg/min) or 50-fold (Table 3.6, D128K + Vif$_{IIIB}$, 0.8 pmol/μg/min) in comparison to an 8-fold (Table 3.7, A3G + Vif$_{HXB2}$, 1.2 pmol/μg/min) or 6-fold (Table 3.7, A3G + Vif$_{IIIB}$, 1.8 pmol/μg/min) decrease in A3G activity. These data suggest that there is a significant amount of Vif binding the 118 nt ssDNA substrate in the presence of D128K, causing a large decrease in specific activity. To investigate this further we determined the apparent K_d of a D128K and Vif$_{HXB2}$ mixture and found the value (Table 3.5, K_d of 72 nM) was almost identical to that of Vif$_{HXB2}$ alone (Table 3.5, K_d of 67 nM). These data support the hypothesis that Vif has a dominant role in ssDNA binding in the presence of D128K that has an approximately 3-fold higher apparent K_d than Vif (Table 3.5, K_d of 183 nM). Similar results were found with a D128K and Vif$_{IIIB}$ mixture binding to the 118 nt ssDNA substrate (Table 3.5, K_d of 51 nM). In contrast, an A3G and Vif$_{HXB2}$ or A3G and Vif$_{IIIB}$ complex had an apparent K_d (Table 3.5, A3G and Vif$_{HXB2}$, K_d of 158 nM; A3G and Vif$_{IIIB}$, K_d of 135 nM) that was above that of Vif$_{HXB2}$ or Vif$_{IIIB}$ alone (Table 3.5, Vif$_{HXB2}$, K_d of 67 nM; Vif$_{IIIB}$, K_d of 79 nM). Based on these data, it appears that Vif molecules that interact with A3G may not be able to bind the ssDNA substrate efficiently through the Vif DNA binding domain since it is near the A3G interaction residues. This would lower the effective DNA binding concentration of Vif and result in an increase in the apparent K_d value (Table 3.5). However, the Vif variants may still contact the ssDNA since the A3G/Vif complex apparent K_d values are slightly below that of A3G alone (Table 3.5, K_d of 203 nM).

In contrast to A3G, which has a decrease in processivity in the presence of Vif (Fig. 3.4),

we do not observe a decrease in D128K processivity on the 118 nt substrate (data not shown) or in the HIV-1 replication assay (Fig. 3.6 and Fig. 3.7) in the presence of the Vif variants, despite the decrease in specific activity (Table 3.7). Together with binding data (Table 3.5), the results indicate that Vif can bind ssDNA and compete with D128K for substrate access, which results in a decrease in D128K specific activity (Table 3.7). This activity is lessened in the presence of A3G (Table 3.7) which binds Vif with high affinity (Table 3.5). Therefore, the data support a model where the inhibition of processivity and specific activity by Vif are two different functions mediated by an A3G-Vif or ssDNA-Vif interaction, respectively.

3.4 Discussion

Although Vif induces polyubiquitination and degradation of A3G in HIV-1 infected cells [70, 225, 246], the clearance of A3G from the host cell may not be complete [93, 124]. Vif can use alternate ways of inhibiting A3G activity such as suppressing A3G mRNA translation, blocking virus encapsidation of A3G, and inhibiting A3G deamination activity [64, 223, 298, 299, 348, 349]. Here we examined the mechanism by which Vif can inhibit A3G deamination activity. Our data demonstrate that Vif changes the processive scanning mechanism of A3G (Fig. 3.4) and this appears to account for the observed decrease in A3G-induced mutagenesis occurring in a model HIV-1 replication system (Fig. 3.2). The data indicate that the change in A3G processivity by the Vif variants is due to a protein-protein interaction since A3G D128K-induced mutagenesis is unaffected by Vif (Fig. 3.6). The implications of the data are that any residual A3G that escapes Vif-mediated degradation and becomes encapsidated into HIV-1 virions in the presence of Vif is likely to be ineffective at restricting HIV-1 replication [93, 124]. Vif and A3G co-encapsidation could also be induced by a therapeutic option that prevents Vif interaction with ubiquitin ligase components, but retains Vif interaction with A3G [311]. Of note, while this work was being completed it was discovered that CBF-β stabilizes Vif in cells [68, 69]. We confirmed that a CBF-β/Vif complex similarly affects A3G processivity and specific activity as the GST-Vif used in our study (Fig. 3.8, Table 3.8, and see Material and Methods), which provides additional evidence that Vif is likely to affect A3G deamination ability *in vivo*. The results of the model HIV-1 replication assay for A3G (Fig. 3.2 and Fig. 3.5) and D128K (Fig. 3.6 and Fig. 3.7) in the presence of the Vif variants supports a model in which a Vif-A3G complex results in a decreased ability of A3G to deaminate C→U processively, resulting in a decrease in mutagenesis. However, we examined if there were additional effects caused by Vif binding
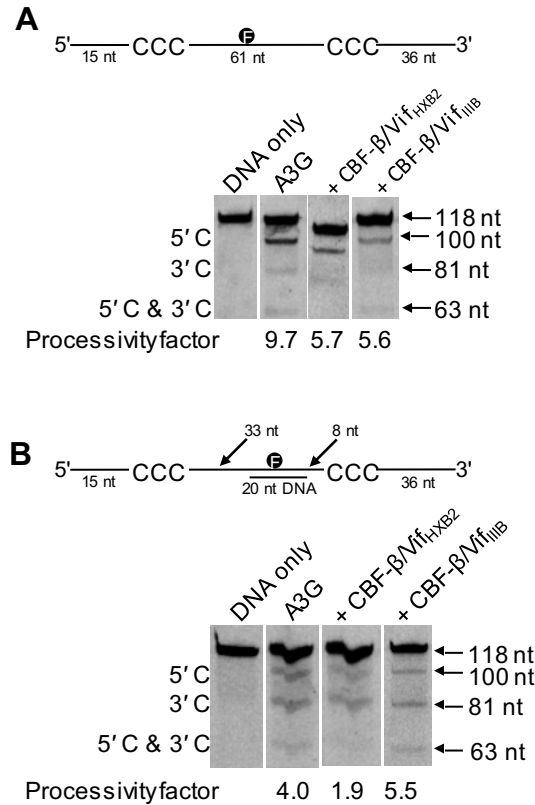
Figure 3.8 Determination of processivity of A3G in the absence and presence of CBFβ/Vif$_{HXB2}$ and CBFβ/Vif$_{IIIB}$. CBFβ/Vif$_{HXB2}$ and CBFβ/Vif$_{IIIB}$ demonstrate a similar effect on A3G as GST-Vif$_{HXB2}$ and GST-Vif$_{IIIB}$ (Fig.3, A-B). A, deamination of a 118 nt F-labeled ssDNA substrate by A3G. Two CCC motifs within the ssDNA sequence are spaced 61 nt apart. Single deaminations of the 5'C and 3'C are detected as the appearance of labeled 100- and 81 nt fragments, respectively; double deamination of both C residues on the same molecule results in a 63 nt labeled fragment (5'C & 3'C). A3G processivity is decreased by ~1.7-fold in the presence of CBF-β/Vif$_{HXB2}$ (+ CBF-β/Vif$_{HXB2}$) and CBF-β/Vif$_{IIIB}$ (+ CBF-β/Vif$_{IIIB}$). B, deamination of the substrate described for (A), but with a 20 nt ssDNA annealed between the two CCC motifs which can block the sliding component of A3G processivity. This results in a ~2-fold decrease in A3G processivity. Processive deaminations are still observed due to the jumping component of A3G processive movement. In the presence of CBF-β/Vif$_{HXB2}$ the processivity of A3G is nearly absent indicating that the jumping component of A3G has been inhibited. This is in contrast to the Vif$_{IIIB}$ that has no effect on A3G jumping movement. The measurements of processivity (Processivity factor) are shown below the gel. The standard deviation for the processivity factors are (A) A3G, 1.4; + CBFβ/Vif$_{HXB2}$, 0.6; + CBFβ/Vif$_{IIIB}$, 0.3 (B) A3G, 0.8; + CBFβ/Vif$_{HXB2}$, 0.9; + CBFβ/Vif$_{IIIB}$, 0.7.

Table 3.8 Specific activities of A3G the absence and presence of CBF-β/Vif variants.

Reaction condition	Specific Activity (pmol/μg/min)
A3G	8.0 ± 0.8
A3G + CBF-β/Vif$_{HXB2}$	1.2 ± 0.5
A3G + CBF-β/Vif$_{IIIB}$	1.3 ± 0.7

The specific activity was determined using the 118 nt substrate and the values are shown with the standard error of the mean that was calculated from three independent experiments. The specific activities of A3G the absence and presence of CBF-β/Vif variants are similar to those obtained with the GST-Vif variants (Table 3.7).

ssDNA. We reasoned that the ssDNA binding effects of Vif would manifest as a decrease in specific activity of A3G or D128K due to competition for substrate binding. This cannot be observed in the HIV-1 replication assay due to substrate access being limited by RNA degradation and (+)DNA synthesis. While this alone argues that the inhibition of A3G processivity by Vif is of primary importance, we examined the effect of Vif binding a synthetic ssDNA substrate to fully characterize the mechanisms by which Vif can affect A3G deamination activity. We hypothesized that because the residues of Vif that interact with A3G are near the Vif RNA/DNA binding domain [355], the A3G-Vif complex may contact ssDNA more efficiently through A3G DNA binding residues. This would argue that there is a minimal effect of Vif on A3G specific activity. Consistent with the model is the observed smaller decrease in A3G specific activity (Table 3.7, 6- to 8-fold) than D128K specific activity (Table 3.7, 50- to 80-fold) in the presence of the Vif variants. As such, our data are in agreement with previously published literature that Vif can decrease A3G deamination activity [298, 299], but here we find that the mechanism by which this occurs is Vif affecting primarily the processive scanning (Fig. 3.4, *A-C* and Fig. 3.5) rather than the specific activity (Table 3.7) of A3G. This is also supported by rotational anisotropy data where Vif variants bind ssDNA with the same apparent K_d in the absence or presence of D128K, but have an increase in the apparent K_d in the presence of A3G (Table 3.5). Based on the data obtained with D128K where we observed a decrease in specific activity (Table 3.7), but not processivity (Fig. 3.6 and Fig. 3.7), it appears that these two inhibitory functions are implemented by free Vif and A3G bound Vif, respectively (Table 3.5 and 3.7). The Vif-bound A3G does not efficiently induce mutagenesis of (-)DNA suggesting that both processive jumping and sliding movements are required, in agreement with other reports [133, 139].

Vif is known to physically interact with A3G and Asp^{128} is a key residue that mediates this interaction [280]. The Asp^{128} is predicted to be located on loop 7 of the N-terminal half of A3G (Fig. 3.1). The predicted loop 7 is functionally important because the amino acids $^{124}YYFW^{127}$ located in this region mediate virion incorporation [137, 138], oligomerization [138, 140] and the jumping component of A3G processivity [139]. The proximity of Asp^{128} to the key processivity residues, $^{126}FW^{127}$, which mediate A3G jumping movements [139], provides a logical explanation of how Vif_{HXB2} can block the jumping ability of A3G (Fig. 3.4, A-C).

Together with the high affinity of A3G for binding Vif_{HXB2} (Table 3.5), the data support that Vif_{HXB2} inhibition of A3G deamination ability is primarily through a protein-protein interaction. It has been demonstrated that a loss of jumping ability in the A3G F126A/W127A mutant resulted in a large decrease of A3G-induced mutations in the HIV-1 replication assay [139] and we observe the same phenomenon here (Fig. 3.2, *A-B* and *D-E*). The decreased mutagenesis observed in the HIV-1 replication assay is likely to correlate with a decrease in HIV-1 infectivity since virion encapsidated F126L or W127A mutants are less effective at restricting HIV-1 in single cycle infectivity assays [137]. We have hypothesized that this is because A3G requires the jumping component of its processive mechanism to effectively transverse the RNA/DNA hybrid regions left by RT RNase H activity [139]. In the presence of Vif_{HXB2}, A3G is capable of sliding, but any surrounding RNA/DNA hybrid regions would essentially lock the A3G into an isolated area of (-)DNA. Due to the extended residence time A3G has on ssDNA, a half-life of ~4 min [141], remaining in a confined (-)DNA region would result in an ineffective search over the whole (-)DNA (Fig. 3.2, compare A and B and D and E).

Interestingly, Vif_{IIIB} has a different effect on A3G processivity than Vif_{HXB2} and inhibits A3G sliding (Fig. 3.2, A-C, Fig. 3.4). Based on phylogenic analysis, Vif_{IIIB} is less similar to Vif_{HXB2} than other variants, such as Vif_{ELI-1} or Vif_{YU-2} (Fig. 3.9) providing a rationale for the difference in their effect on A3G and suggests that the inhibition of A3G jumping rather than sliding is more likely to be observed in other variants. It has been suggested that the sliding component of A3G is mediated by helix 6 [139]. Helix 6 is predicted to be adjacent to loop 7 (Fig. 3.1). The contacts of Vif with A3G have been shown to encompass $^{128}DPD^{130}$ on loop 7 and hypothesized to also extend beyond loop 7 [267], which is supported by a mutagenesis study [356]. However, the A3G-Vif interface identified by Letko *et al*. [284] through functional assays and molecular docking did not recognize helix 6 to be involved in A3G-Vif interaction. This may explain why the addition of Vif_{IIIB} does not cause a more severe effect on A3G than Vif_{HXB2} in the model HIV replication system (Table 3.1). Vif_{IIIB} may stabilize itself on A3G primarily by interacting with loop 7 and subsequently establish an unstable interaction with helix 6. Conformational change and structural flexibility in A3G could be induced upon Vif_{IIIB} binding to A3G (presumably through loop 7) [312], which allows Vif_{IIIB} to contact helix 6 and affect the sliding motion of A3G. Nevertheless, random mutagenesis identified residues in the N-terminal domain predicted helices 3, 4, and 6 and predicted loops 1 and 3 that interact with Vif_{IIIB} [356].
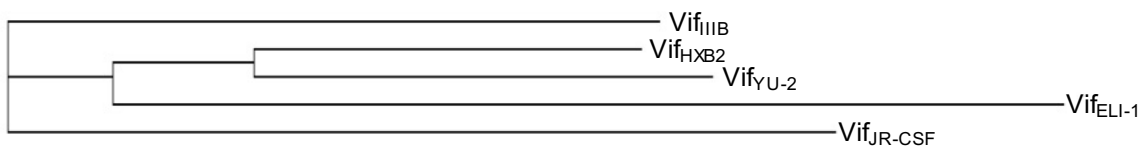
Figure 3.9 Phylogram of Vif variants. The GenBank accession numbers for the Vif variants are: M19921 (Vif$_{IIIB}$), K03455 (Vif$_{HXB2}$), K03454 (Vif$_{ELI-1}$), M38429 (Vif$_{JR-CSF}$), and AY569174 (Vif$_{YU-2}$). Phylogram was made using Clustal W.

This may be due to a charged surface on A3G that interacts with Vif through multiple structural components. This appears to be the case for A3F, which interacts with Vif through its C-terminal half, where a negatively charged surface surrounding predicted C-terminal helices 2, 3, and 4 may interact with Vif in addition to a key residue, Glu 289 [205, 285, 287, 288]. These reports provide a supporting rationale for how A3G sliding, which is mediated by helix 6 [139], could be affected in our study by Vif_{IIIB}. Furthermore, Vif_{IIIB} and Vif_{HXB2} have been found to differentially mediate the extent of APOBEC3 enzyme degradation [295], which implies they may interact with A3G differently beyond residues $^{128}DPD^{130}$. Since A3G and Vif_{IIIB} interact with high affinity (Table 3.5, K_d of 78 nM), the data support the conclusion that Vif_{IIIB} inhibition of A3G activity is primarily through a protein-protein interaction that disrupts the sliding component of the processive ssDNA scanning mechanism.

It has previously been reported that any type of small molecules that could be designed to bind A3G and block its interaction with Vif may not be a relevant therapeutic option if A3G functions imparted by the residues adjacent to Asp^{128}, such as virion encapsidation, are also affected [267, 300]. For this reason, alternate therapeutic options have been put forward such as small molecules that bind Vif [301] or the ubiquitin ligase complex [309, 311] to inhibit degradation of A3G. The latter therapeutic option would presumably allow a maintained interaction between Vif and A3G. Based on our data, a therapeutic option that left Vif in the cell allowing it to form a Vif-A3G complex would not enable virion encapsidated Vif-A3G to reliably restrict HIV-1 replication (Fig. 3.2, *A-F* and Fig. 3.3). Because the number of A3G-induced mutations has been correlated with HIV-1 inactivation [124, 151], an implication of our results is that the Vif-A3G complex could lead to sublethal mutagenesis of HIV-1. However, reports published recently suggest that A3G-induced sublethal mutagenesis is likely to be an rare event and it does not promote HIV-1 evolution more than RT replication errors [75, 76], which indicates A3G can act in the host's favor and irretrievably inactivate HIV-1 through lethal mutagenesis, regardless of the presence of Vif inside virions. Nevertheless, our data indicate that Vif can make virion encapsidated A3G less effective and small molecular inhibitors that disrupt Vif-mediated degradation pathway of A3G thus allowing the co-packaging of Vif-A3G could be provide a potential avenue for therapeutic intervention.

Table 3.9 Primers and DNA substrates.

Name	Sequence
Vif$_{IIIB}$ Forward	ATG GAG AAC CGG TGG CAG GTG ATG
Vif$_{IIIB}$ Reverse	CTA GTG TCC ATT CAT TGT ATG GCT CCC
Vif$_{HXB2}$ Forward	ATG GAA AAC AGA TGG CAG GTG
Vif$_{HXB2}$ Reverse	C TA GTG TCC ATT CAT TGT GTG
CBF-β Forward	ATG CCG CGC GTC GTG CCC GAC CAG A
CBF-β Reverse	CTA GGG TCT TGT TGT CTT CTT GCC AGT TAC TGC CAG
D128K SDM Forward	CC CGC CTC TAC TAC TTC TGG AAG CCA GAT TAC CAG GAG GCG CTT C
D128K SDM Reverse	GA AGC GCC TCC TGG TAA TCT GGC TTC CAG AAG TAG TAG AGG CGG G
L60 (deamination assay and anisotropy)	GAA TAT ATG TTG AGA CCC AAA GTA ATG AGA GAT TGA [dT-Fam] TAG ATG AGT GTA ATG TGA TAT ATG TGT ATG AAA GAT ATA AGA CCC AAA GAG TAA AGT TGT TAA TGT GTG TAG ATA TGT TAA
L60 complementary DNA	CTT TCA TAC ACA TAT ATC AC
L60 complementary RNA	CUU UCA UAC ACA UAU AUC AC
Protease (-) DNA (anisotropy)	[5'-Fam] GAT AAA ACC TCC AAT TCC CCC TAT CAT TTT TGG TTT CCA TTT CCC TGG CAA ATT TAT ATC TTC TAA TAC TGT ATC ATC TGC TCC TGT ATC TAA TAG AGC TTC TTT CAG TTG TCC TCC TAT

Abbreviations are: SDM, site directed mutagenesis; Fam, fluorescein.

3.5 Material and Methods

3.5.1 Protein Expression and Purification. Recombinant baculovirus production for expression of GST-A3G, GST-A3G D128K, GST-Vif$_{IIIB}$, GST-Vif$_{HXB2}$ or GST-nucleocapsid protein (NC) in *Sf*9 cells was carried out using the transfer vector pAcG2T (BD Biosciences), as previously described [132, 139]. Vif$_{NL4-3}$ was cloned from pcDNA-hVif vector that encodes the Vif protein of HIV-1 NL4-3, which was obtained from the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH from Dr. Stephan Bour and Dr. Klaus Strebel [357]. HIV-1 NL4-3 and HIV-1 IIIB have the same Vif sequence and we have referred to Vif from this clone as Vif$_{IIIB}$. Vif$_{HXB2}$ was cloned from pHIV-*gpt*(wt), a gift from Dr. David Kabat (Oregon Health and Science University). Site directed mutagenesis was used to create the A3G D128K clone. Cloning primers for Vif variants and the site directed mutagenesis primers were obtained from Integrated DNA Technologies and are listed Table 3.9. *Sf*9 cells were infected with recombinant virus at a multiplicity of infection (MOI) of 1, except for GST-Vif$_{IIIB}$ and GST-Vif$_{HXB2}$ which were infected at an MOI of 4 or 2, respectively. Recombinant baculovirus infected *Sf*9 cells were harvested after 72 h of infection, except for GST-Vif$_{IIIB}$ and GST-Vif$_{HXB2}$ which were harvested after 48 h of infection. Cells were lysed in the presence of RNase A and the proteins (A3G, A3G D128K, and NC) were purified as described previously [140] to obtain protein that was cleaved from the GST tag and 95% pure. The Vif variants were eluted from the glutathione-sepharose resin (GE Healthcare), as previously described [132], with the GST tag to maintain their stability. Eluted GST-Vif variants were dialysed against 100 mM Tris pH 7.5, 250 mM NaCl, 10% glycerol, and 1 mM DTT and are 85% pure. Protein fractions were stored at -80 °C. HIV-1 reverse transcriptase p66/p51[358] was generously provided by Dr. Stuart F.J. Le Grice (NCI, National Institutes of Health).

Recombinant expression of Vif$_{HXB2}$ or Vif$_{IIIB}$ in the presence of CBF-β was done in *E. coli* using the double expression vector, pET-DUET (Novagen). CBF-β tagged was hexa histidine tagged at the N-terminus. Vif$_{IIIB}$ was cloned from pcDNA-hVif vector that encodes the Vif protein of HIV-1 NL4-3, which was obtained from the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH from Dr. Stephan Bour and Dr. Klaus Strebel [357]. Vif$_{HXB2}$ was cloned from pHIV-*gpt*(wt), a gift from Dr. David Kabat (Oregon Health and Science University). CBF-β isoform 2 (Accession number BC018509) was obtained from Open Biosystems. Cloning primers were obtained from Integrated DNA Technologies and are listed in

Table 3.9. Induction of exponential phase BL21(DE3) *E. coli* cells with IPTG (0.2 mM) proceeded at 16 °C for 18 h. Cells were lysed in lysis buffer containing 20 mM Tris pH 7.5, 150 mM NaCl, 1 mM DTT, complete protease inhibitor (Roche), 1 mg/mL lysozyme, and 40 μg/mL RNAseA in combination with sonication. Cleared lysates were applied to HisTALON beads (Clontech) and purification of the protein was done following manufacturer's instructions. Eluted CBF-β/Vif was applied to a HiLoad Superdex 75 (16/600) gel filtration column (GE Healthcare) and dimer fractions eluted in running buffer (100 mM Tris pH 7.5, 150 mM NaCl, and 1 mM DTT) were adjusted to contain 10% glycerol before storage at -80 °C. Protein preparations are >95% pure.

3.5.2 Model HIV-1 Replication Assay. The model HIV-1 replication assay, which measures A3G-induced mutagenesis of ssDNA during reverse transcription of an RNA template, was performed as described previously [139]. A synthetic (+) RNA genome is synthesized *in vitro* that contains a PPT, 120 nt of the catalytic domain of the HIV-1 *protease* (*prot*), and *lacZα* (248 nt). The PPT enabled second strand ((+)DNA) synthesis to occur. The *lacZα* was used as a reporter gene for mutations by blue/white screening. The originating sequence of the HIV-1 protease gene is from a clone obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: p93TH253.3 from Dr. Feng Gao and Dr. Beatrice Hahn[359]. The RNA template (25 nM) was annealed to a 24 nt DNA primer and incubated with NC (1.5 μM), RT (1.2 μM), and dNTPs (500 μM) in RT buffer (50 mM Tris pH 7.4, 40 mM KCl, 10 mM MgCl$_2$, 1 mM DTT) in the presence or absence of 100 nM of A3G or A3G D128K. A3G, RT and NC are added to the reaction at ratios to the RNA that are estimated to be found in virions [5, 92, 360, 361]. Reactions containing A3G or A3G D128K were conducted in the presence or absence of 175 nM Vif$_{IIIB}$ or Vif$_{HXB2}$. Synthesized DNA was PCR amplified using Pfu Turbo C$_x$ (Agilent Technologies) and cloned into a pET-Blue vector backbone that would allow the experimentally synthesized *lacZα* to be used for α-complementation[139]. Twenty-five mutated clones for each condition tested were analyzed. DNA sequencing was carried out at the National Research Council of Canada (Saskatoon, Saskatchewan). A t-test was used for statistical analysis of sequences.

3.5.3 Deamination Assays. The 118 nt ssDNA substrate was obtained from Tri-Link Biotechnologies and the sequence is Table 3.9. A3G and A3G D128K (10-20 nM) were incubated with 100 nM fluorescein (F)-labeled ssDNA and 700 nM Vif (IIIB or HXB2) in RT

buffer. The substrate to A3G ratio is kept high to ensure reactions were carried out under single-hit conditions, i.e., < 15% substrate consumed, where a single A3G could interact with a particular ssDNA substrate at most once[362]. Reactions were incubated at 37°C for 2.5 - 40 min. Deaminations were detected by resolving DNA that had been treated with Uracil DNA Glycosylase and heated under alkaline conditions on a 10% v/v denaturing polyacrylamide gel, as described previously[132]. Gel pictures were obtained using a Typhoon Trio (GE Healthcare) multipurpose scanner and analysis of integrated gel band intensities used ImageQuant software (GE Healthcare). Under these conditions, a processivity factor can be determined by comparing the total number of deaminations occurring at two sites on the same DNA substrate to a calculated theoretical value of the expected deaminations that would occur at those two sites if the events were uncorrelated, i.e., not processive [139]. The specific activity was calculated from single-hit condition reactions by determining the picomoles of substrate consumed per minute for a microgram of enzyme. The addition of DNA was used to start the deamination reactions. Prior to the addition of DNA the reaction components were prewarmed to 37°C.

3.5.4 Steady State Rotational Anisotropy Assays. Steady state fluorescence depolarization (rotational anisotropy) was used to measure protein-nucleic acid and protein-protein binding affinities using a F-labelled binding partner. A QuantaMaster QM-4 spectrofluorometer (Photon Technology International) with a dual emission channel was used to collect data and calculate anisotropy. Measurements were made at 21°C. Samples were excited with vertically polarized light at 495 nm (6 nm band pass) and vertical and horizontal emissions were measured at 520 nm (6 nm band pass). Apparent dissociation constants (K_d) were obtained by fitting to a rectangular hyperbola or sigmoidal curve using Sigma Plot 11.2 software.

A3G, A3G D128K, Vif_{IIIB}, and Vif_{HXB2} were tested for their ability to bind a 118 nt ssDNA substrate, which is listed in Table 3.9 and described in the text. Reactions were 50 μL and contained F-labelled ssDNA (50 nM) in RT buffer and A3G (0 - 700 nM) or Vif (0 - 200 nM) were titrated into the reaction. Alternatively, A3G or D128K were mixed with a Vif variant at an equimolar ratio before titration (0-350 nM) in to the reaction.

For protein-protein binding, A3G or A3G D128K were labeled with fluorescein using the Fluorescein-EX Protein Labeling Kit (Life Technologies). The F-A3G or F-D128K (50 nM) was used in a reaction mixture (50 μL) which contained RT buffer and a titration of Vif_{IIIB} or Vif_{HXB2} (0 - 200 nM).

3.6 Acknowledgements

4.0 Investigation of the degradation-independent and-dependent functions of Vif against A3s

Two landmark papers published in late 2012 reported the core binding factor beta (CBFβ) acts as a pivotal Vif regulator by being part of the E3 ubiquitin ligase complex that is involved in polyubiquitinating and degrading Vif [68, 69] and these findings have prompted many questions to be revisited, including the physiological form of Vif that participates in the non-degradative Vif-mediated inhibition of A3 proteins. Understanding the physiologically active form of Vif allows the development of more relevant model system to study Vif-A3 interactions *in vitro*. Vif needs to bind HIV-1 5'UTR to achieve virion encapsidation [363] and we have found that both forms of Vif, namely Vif/CBFβ heterodimer and Vif/CBFβ/EloB/C heterotetramer, are able to bind 5'UTR with high affinity, as the apparent K_d falls into the nanomolar range (Table 4.1), albeit some Vif variants-specific differences in binding affinities were observed. Whether the differences in RNA binding leads to different virion encapsidation efficiencies of the Vif forms was tested by examining the form of Vif encapsidated into HIV virions using western blot and antibodies against each native protein. The preliminary results indicate that Vif encapsidates in virions without CBFβ and EloB/C (Fig. 4.0). It may be that RNA binding stabilizes Vif and displaces CBFβ and the EloB/C complex.

Through the investigation of the degradation-independent inhibition of A3G by Vif (see Chapter 3) [183], we used A3G D128K to help dissecting whether Vif inhibits A3G's deamination activity through a protein-protein or protein-DNA interaction. We observed that upon mutating ^{128}D to an alanine, instead of lysine, Vif could still cause a decrease in D128A-induced mutagenesis in our reconstituted HIV replication system (Table 4.2), in contrast to what was observed for D128K [183] (Fig. 3.6). This data has made us realize that not just amino acids contacts, but the overall charge of the Vif-A3G interface was of importance. In addition, due to Vif_{HXB2} and Vif_{IIIB} influence A3G DNA scanning and processivity differently (Fig. 3.4), we have obtained evidence that different Vif variants could use different amino acid contacts on A3G beyond the key ^{128}D amino acid. Based on these findings, we next set out to test whether the different A3G-Vif interaction interface would have an impact on the strength of the interaction and if this would influence the efficiency of Vif-induced degradation.

Table 4.1 Binding affinities of different Vif forms for the HIV-1 5'UTR

Vif form	Apparent binding constant (K_d), nM
Vif_{HXB2}/CBFβ heterodimer	16 ± 1
Vif_{IIIB}/CBFβ heterodimer	125 ± 18
Vif_{HXB2}/CBFβ/EloB/C heterotetramer	146 ± 20
Vif_{IIIB}/CBFβ/EloB/C heterotetramer	61 ± 9

Table 4.2 D128A-mediated mutation frequencies in a model HIV replication system in the absence and presence of Vif_{HXB2} and Vif_{IIIB}.

Reaction condition	Population mutation frequency
D128A	0.90
D128A + Vif_{HXB2}	0.66[b]
D128A + Vif_{IIIB}	0.82[c]

The population mutation frequency is defined as the ratio of white colonies to total colonies.
[a] Significant difference was designed as $p \leq 0.001$ versus A3G values.
[b] Significant difference was designed as $p \leq 0.01$ versus A3G values.
[c] Significant difference was designed as $p \leq 0.05$ versus A3G values.
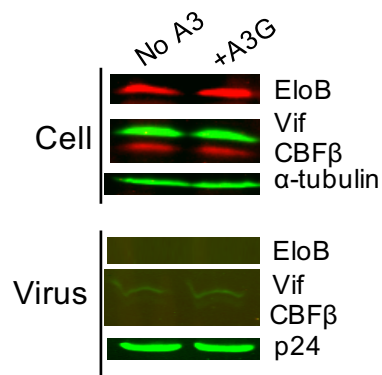
Figure 4.0 Form of Vif in virions. 293T cells infected with HIV was produced in the absence or presence of HA-tagged A3G and levels of Elongin B (EloB), Vif, and CBFβ expressed in cells and encapsidated into HIV virions were determined using immunoblotting. The expression of EloB, Vif and CBFβ were detected using the native antibody. The loading control for cell lysates was α-tubulin and for virions was p24. Despite the Vif/CBFβ/EloB/C heterotetramer being required for Vif function in cells, in virions we could only detect Vif alone. No Elo B was detected in virions. We did not have an appropriate antibody to detect Elongin C (EloC); however, that the EloB/C form an obligate heterodimer suggesting that neither is present in the virion with Vif. CBFβ was also not present. Thus, the stable form of Vif in cells and virions may be different. In virions, RNA binding may stabilize Vif.

All experiments listed in chapter 5 were performed by Y.F.. T.T.B., Y.F. and L.C. conceived and designed the experiments. T.T.B, Y.F. and L.C. analyzed the data. T.T.B and L.C. wrote the article. T.T.B, Y.F. and L.C. revised the article.

5.0 Determinants of efficient degradation of A3 restriction factors by HIV-1 Vif

Tayyba T. Baig[1], Yuqing Feng[1], Linda Chelico[1]
[1]Department of Microbiology & Immunology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada, S7N 5E5

Running title: Determinants of Vif-mediated degradation.

The information in this chapter was previously published:

Baig, T.T., Feng, Y., and Chelico, L. (2014). Determinants of efficient degradation of APOBEC3 restriction factors by HIV-1 Vif. *Journal of Virology*. 88, 14380-14395.

5.1 Abstract

The APOBEC3 deoxycytidine deaminases can restrict the replication of HIV-1 in cell culture to differing degrees. The effects of APOBEC3 enzymes are largely suppressed by HIV-1 Vif that interacts with host proteins to form a Cullin5-Ring E3 ubiquitin ligase that induces [48]K-linked polyubiquitination (poly-Ub) and proteasomal degradation of APOBEC3 enzymes. Vif variants have differing abilities to induce degradation of APOBEC3 enzymes, and the underlying biochemical mechanisms for these differences is not fully understood. We hypothesized that by characterizing the interaction of multiple APOBEC3 enzymes and Vif variants we could identify common features that resulted in Vif-mediated degradation and further define the determinants required for efficient Vif-mediated degradation of APOBEC3 enzymes. We used Vifs from HIV-1 IIIB and HXB2 to characterize their induced degradation of and interaction with APOBEC3G and APOBEC3H in 293T cells. We quantified the APOBEC3G-Vif and APOBEC3H-Vif interaction strengths *in vitro* using rotational anisotropy. Our biochemical and cellular analyses of the interactions support a model in which the degradation efficiency of Vif_{IIIB} and Vif_{HXB2} correlated with both the binding strength of the APOBEC3-Vif interaction and the APOBEC3-Vif interface, which differs for APOBEC3G and APOBEC3H.

5.2 Introduction

The APOBEC3 (A3) deoxycytidine deaminases can act as intracellular restriction factors against replication of HIV-1 (referred to as HIV) [99]. A3 enzymes that are encapsidated into budding HIV virions can restrict HIV replication in the next target cell by deaminating cytosine in (-) strand single-stranded DNA, which forms mutagenic uracils and results in numerous C/G→T/A transition mutations that can inactivate the virus [22-24]. In CD4[+] T cells it appears that four of the seven A3 members, A3D, A3F, A3G, and A3H Haplotype II (referred to as A3H) are primarily responsible for HIV restriction [78]. Nonetheless, HIV can successfully infect cells where A3 enzymes are highly expressed due to the viral infectivity factor (Vif) protein [246, 346]. Vif acts as a substrate receptor for a Cullin5-Ring E3 (Cul5-E3) ubiquitin ligase complex, which can induce polyubiquitination (poly-Ub) and degradation of A3 enzymes [70, 248]. This process is mediated by Vif binding to host Cullin5 and the Elongin B/C heterodimer (EloB/C) through specific motifs in Vif that mimic human SOCS2 [72, 91, 249-251]. Vif also binds with the transcription cofactor CBFβ for thermodynamic stability [68, 69]. Within this E3 ligase complex, Rbx2 recruits an E2 ubiquitin conjugating enzyme to induce [48]K-linked poly-Ub of A3

enzymes, which is concomitant with their proteasomal degradation [63, 68, 70, 225, 247]. Vif interacts with A3s through its N-terminal domain (NTD) in distinct regions for A3G ([40]YRHHY[44]), A3H ([39]F, [48]H), or A3C, A3D and A3F ([14]DRMR[17]) in conjunction with secondary binding sites [240, 265, 269-273, 364]. The positively charged surfaces of Vif interact with negatively charged amino acids on the A3 enzyme, with some contribution from hydrophobic amino acid interactions, depending on the interface [265]. There are three distinct structural motifs on the A3 enzymes that interact with a corresponding Vif region. These can be categorized in three classes: A3G-, A3H-, or A3C/A3F/A3D- like [265].

A3G is a double Z-domain enzyme that primarily interacts with Vif through residues [128]DPD[130] on predicted loop 7 in the NTD, [128]DPD[130] [267]. The [128]D amino acid was identified as essential for Vif-mediated degradation and mediation of a cross-species barrier to simian immunodeficiency virus (SIV) by mutating [128]D to [128]K, as found in African Green Monkey or Rhesus Macaque A3G [278-281]. Multiple research groups found that Vif did not interact or induce degradation of the A3G D128K mutant. However, if the [128]D was mutated to an alanine, Vif-mediated degradation occurred, demonstrating that the overall charge of the A3G-Vif interface was of importance and not only the amino acid *per se* [278, 280, 281]. The overall interface of A3G that interacts with Vif was studied through a MAPPIT analysis and it was found that Vif[IIIB] could interact not only with [128]D located on A3G structural loop 7 but also other adjacent structures, such as helix 6 [283]. The MAPPIT analysis was in agreement with other studies that identified larger regions of Vif binding [131, 175, 183, 284]. Further, a study by our group found that Vif[IIIB] and Vif[HXB2] interacted with different regions of the A3G NTD beyond [128]D [183]. This enabled each Vif variant to cause distinct changes in A3G enzymatic activity [183].

For A3H, Vif variant differences in the ability to induce A3H degradation are widespread and have been extensively studied [240, 276]. Data from cell culture and HIV-1 infected patients demonstrates that amino acids 39 and 48 in Vif are important for both the ability of Vif to induce degradation of and interact with A3H [240]. For example, Vif[LAI] which contains [48]H can induce degradation, but not Vif[IIIB] that contains a [48]N. It was concluded that Vif[IIIB] is unable to degrade A3H due to an inability to efficiently associate with the enzyme [276]. For A3H, the interface that interacts with Vif occurs on predicted helix 4 with [121]D being the primary determinant in Vif-mediated degradation efficiency [195, 211, 265, 290, 292]. Zhen *et al*. found that a D121K

mutant was unable to interact with Vif and concluded that similar to the example of A3G D128K a lack of Vif-induced degradation correlates with a lack of a A3-Vif interaction [290].

Vif interacts with the A3F/A3D interface through A3 helices 2, 3, 4, and β-strand 4 [205, 285, 287, 288, 365, 366]. For A3F and A3D that contain two Z-domains, the structural motif is at their C-terminal domain (CTD). Extensive mutagenesis studies by several groups on A3F demonstrated that mutation of conserved residues on these helices cause less or abolish Vif-mediated A3F degradation [205, 285, 288]. Mutation of A3F [263]L, [264]S or [324]E that are involved in Vif-mediated degradation [205] did not decrease the binding affinity of Vif for A3F [365, 367], suggesting that for A3F the larger Vif binding interface is not as easily disrupted as for A3G and A3H [366]. However, Kitamura *et al.* and Richards *et al.* found that an A3F E324K mutant had a partial decrease in the A3F-Vif interaction [205, 287] and Siu *et al.* found no interaction of an E324Q mutant with Vif [365], demonstrating that there is experimental variability between labs and the relationship between a lack of Vif-mediated degradation and the A3-Vif interaction still requires further investigation.

Despite a large amount of data regarding A3 and Vif interactions, the biochemical basis of how Vif induces A3 substrate poly-Ub and degradation remains to be determined. Studies which examined differences in Vif variants and found specific nonconserved amino acids that were responsible for the ability to induce degradation suggest that the interaction strength with Vif is a primary determining factor in successful Vif-mediated degradation [240, 276]. However, these studies used qualitative co-immunoprecipitations to characterize these interactions. Therefore, to better understand the mechanistic role of Vif as the Cul5-E3 ubiquitin ligase substrate receptor we undertook a study to define the parameters for efficient Vif-induced degradation of an A3 enzyme, using both biochemical and cellular experiments. We hypothesized that the interaction strength between the A3 and Vif was not the sole determinant of degradation efficiency and that the A3-Vif interface also plays a role. To test this hypothesis, we used Vif variants from HIV NL4-3 (IIIB) and HXB2 and A3 enzymes that are either efficiently degraded by these Vifs (A3G) or not efficiently degraded by both Vifs (A3G D128K, A3H) [211, 240, 267, 278-281, 295]. We used cellular experiments to determine Vif-induced degradation efficiency and A3-Vif interactions in cells in conjunction with an *in vitro* quantitative method to determine the binding strength of A3G, A3G D128K and A3H with Vif variant heterotetramers (Vif/CBFβ/EloB/C), the most stable form of Vif [244, 252, 368]. Our

biochemical and cellular analyses of the interactions support a model in which the degradation efficiency of Vifs correlated with the both the A3-Vif binding strength and A3-Vif interface.

5.3 Results

5.3.1 Interaction of A3G with HIV-1 Vif$_{IIIB}$ and Vif$_{HXB2}$. A3G degradation is induced by all HIV Vif variants (referred to as Vif) tested in the literature [211, 240, 295, 369], including Vif$_{IIIB}$ and Vif$_{HXB2}$ used in our study (Fig. 5.1). In Baig *et al.* [312], we have demonstrated the almost equivalent abilities of Vif$_{IIIB}$ and Vif$_{HXB2}$ to induce proteasomal degradation of A3G in a degradation assay where A3G and partially codon optimized Vif expression plasmids were cotransfected into 293T cells. However, Vif$_{IIIB}$ could induce slightly less degradation of A3G than Vif$_{HXB2}$ [312]. Evidence of the ability of Vif$_{HXB2}$ and Vif$_{IIIB}$ to use different amino acid contacts on A3G beyond the key ^{128}D amino acid exists from biochemical studies that examined the effect of Vif on A3G DNA scanning and processivity [183]. This biochemical data prompted us to test whether the different A3G-Vif interactions have an impact on the strength of the interaction and if this influenced the efficiency of Vif-induced degradation.

To obtain a quantitative value for A3G interacting with Vif$_{HXB2}$ and Vif$_{IIIB}$ we used rotational anisotropy. Rotational anisotropy measures the apparent change in rotation speed of a fluorescently labeled molecule due to a change in shape and/or molar mass of the molecule upon binding an unlabeled partner [370]. We fluorescently labeled A3G and titrated Vif heterotetramer (Vif/CBFβ/EloB/C) into the binding reaction. The binding of the fluorescently labeled A3G to the Vif heterotetramer will result in a change in rotation speed (anisotropy) until the fluorescently labeled A3G population is saturated with its binding partner. From these measurements an apparent dissociation constant (K_d) can be calculated. Since Vif must be coexpressed with CBFβ for stability [244], we would be unable to specifically label the Vif and therefore chose to fluorescently label A3G. Usually, upon a fluorescently labeled molecule interacting with a binding partner the increase in size of the complex causes the rotation speed to decrease, which results in an increase in rotational anisotropy. However, with A3G we found that there was a decrease in anisotropy, indicative of an increase in rotation speed (Fig. 5.2A). We hypothesized that this was due to Vif disrupting the dimers of A3G [140, 142] and interacting with a monomer of A3G since the key residues for A3G dimerization (^{126}FW^{127}) are adjacent to
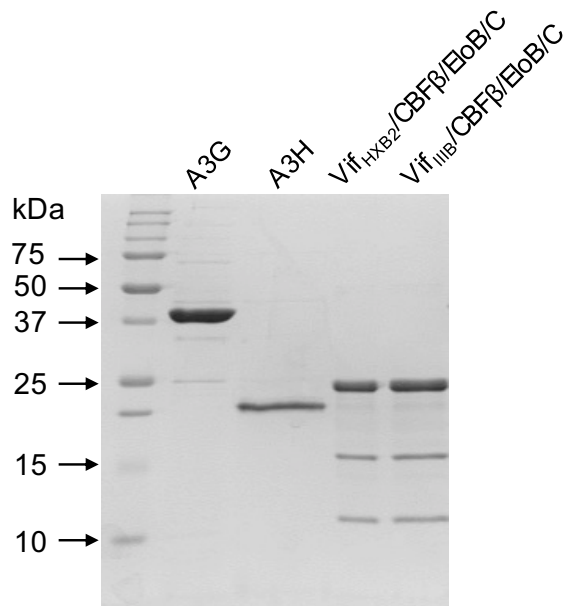
Figure 5.1 Purity of enzymes used for in vitro biochemical assays. The purity of A3G, A3H (haplotype II), Vif$_{HXB2}$/CBFβ/EloB/C, Vif$_{IIIB}$/CBFβ/EloB/C was assessed by resolving 3.5 μg of purified recombinant proteins using 15% SDS-PAGE and staining with Coomassie blue.
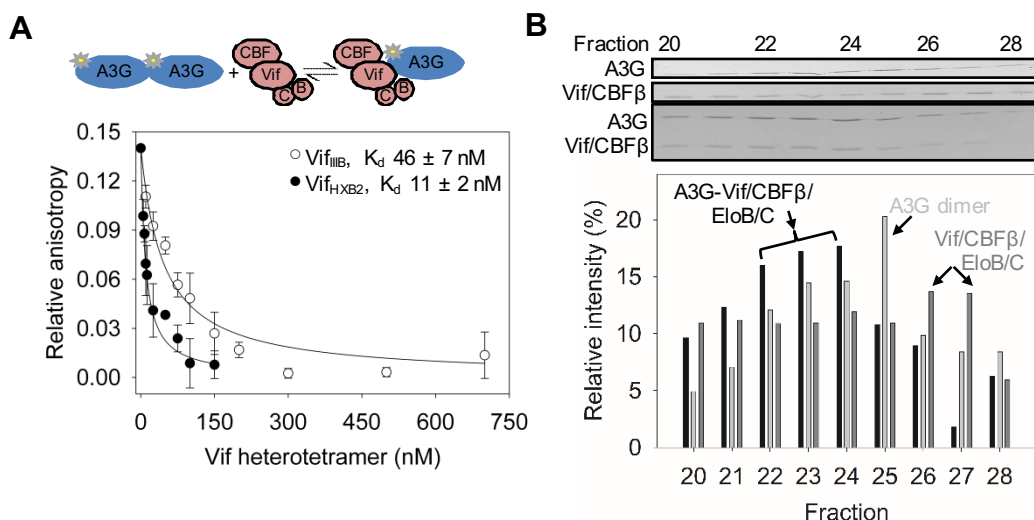
Figure 5.2 Degradation efficiency and interaction of Vif$_{IIIB}$ and Vif$_{HXB2}$ with A3G. (A) Steady state rotational anisotropy (fluorescence depolarization) was used to measure the binding strength between A3G and Vif$_{IIIB}$/CBFβ/EloB/C or Vif$_{HXB2}$/CBFβ/EloB/C (labeled as Vif$_{IIIB}$ or Vif$_{HXB2}$). Anisotropy results were normalized to a start value of 0.14 and end value of zero for comparison. The actual anisotropy start value was between 0.14 and 0.12. The Vif variant heterotetramer was titrated into a binding reaction with fluorescently labeled A3G to determine an apparent dissociation constant (K$_d$). A change in anisotropy indicates an interaction between A3G and the Vif heterotetramer. The decrease in anisotropy indicated that the initial fluorescently labeled complex had decreased in size and/or undergone a conformational change that increased the speed of rotation (see sketch above panel A). The Vif$_{IIIB}$ (apparent K$_d$ 46 ± 7 nM) and Vif$_{HXB2}$ (apparent K$_d$ 11 ± 2 nM) heterotetramers efficiently interacted with A3G. Error bars represent the standard deviation of the mean from three independent experiments. (B) Size exclusion chromatography was used to determine the oligomerization state of an A3G-Vif/CBFβ/EloB/C complex. For comparison, A3G alone and the Vif/CBFβ/EloB/C heterotetramer alone were resolved by size exclusion chromatography. Peak fractions were identified by the UV absorbance of the chromatogram and confirmed by resolving samples by SDS-PAGE. The graph represents the quantification of the coomassie stained bands shown in the SDS-PAGE gels. Comparison to molecular weight standards was used to calculate apparent molecular masses of peak fractions. The Vif/CBFβ/EloB/C heterotetramer resolved as a stoichiometric complex with an apparent molecular mass of 67 kDa (Fractions 26-27). A3G resolved as a dimer with an apparent molecular mass of 74 kDa in the peak fraction (Fraction 25). The A3G-Vif/CBFβ/EloB/C complex resolved in fractions 22-24 which gave an apparent molecular mass range of 157-96 kD. The presence of EloB/C in fractions containing Vif/CBFβ was confirmed by staining with Bio-rad Oriole stain (gel not shown).

[128]D [138, 140]. The existence of this stoichiometric complex (A3G/Vif/CBFβ/EloB/C) was confirmed with size exclusion chromatography (Fig. 5.2B). First, A3G and Vif$_{HXB2}$ heterotetramer were run individually on the size exclusion column. Consistent with previous studies, A3G had an apparent molecular mass of a dimer (Fig. 5.2B, Fraction 25, 74 kDa) and the Vif$_{HXB2}$ heterotetramer was a 1:1:1:1 complex (Fig. 5.2B, Fraction 26, 67 kDa) [140, 258]. The A3G-Vif$_{HXB2}$ heterotetramer complex resolved in three peak fractions with apparent molecular masses ranging from 157– 96 kDa (Fig. 5.2B, Fractions 22-24). If a Vif$_{HXB2}$ heterotetramer were to interact with a monomer of A3G, the expected apparent molecular mass would be 111 kDa, which is within the range of the peak fractions (Fig. 5.2B, Fraction 23-24). Equivalent results were found with Vif$_{IIIB}$ (data not shown). As a result, the apparent dissociation constants obtained were a measure of the ability of the Vif heterotetramer to disrupt A3G oligomerization, which was simultaneous with an interaction between the two proteins. A proportion of the Vif$_{HXB2}$ heterotetramer interacts with a dimer of A3G (Fig. 5.2B, expected apparent molecular mass of 141 kDa, Fraction 22). This is due to the size exclusion chromatography being run with a 1:1 ratio of A3G:Vif heterotetramer to enable resolution of the different species. According to the rotational anisotropy data, we would require a 2-fold excess of Vif heterotetramer to completely disrupt the A3G dimer population in steady-state (Fig. 5.2A, 50 nM A3G saturates at 100 nM Vif heterotetramer).

Consistent with efficient degradation by both Vifs, the apparent K_d of the A3G and Vif heterotetramer complexes were in the low nanomolar range (Fig 5.2A, Vif$_{IIIB}$, apparent K_d 46 ± 7 nM; Vif$_{HXB2}$, apparent K_d 11 ± 2 nM). However, the apparent K_d of Vif$_{IIIB}$ and A3G was approximately 4-fold higher than the apparent K_d of Vif$_{HXB2}$ and A3G (Fig. 5.2A). This difference, albeit only subtle, is consistent with results from the degradation assays where more residual A3G in the presence of Vif$_{IIIB}$ than Vif$_{HXB2}$ was observed (data not shown and [312]) and suggests that small changes in the binding strength influence degradation efficiency. All together these data are consistent with other studies that demonstrated that the determining factor for the ability of Vif to degrade an A3 enzyme is a physical interaction.

5.3.2. HIV-1 Vif$_{IIIB}$ maintains an interaction with A3H in the absence of inducing degradation.
To examine if determinants of the A3-Vif interaction are similar for other A3s, we studied A3H for comparison. A3H has been shown to interact with different amino acids on Vif than A3G and

is differentially sensitive to Vif variants from HIV NL4-3 (IIIB) and LAI [100, 240, 284, 292]. The Vif_{HXB2} used in our study and Vif_{LAI} sequences are identical. Baig *et al.* [312] has shown that A3H was not sensitive to Vif_{IIIB}-mediated degradation, but was sensitive to Vif_{HXB2} degradation and co-immunoprecipitation suggests both Vif_{IIIB} and Vif_{HXB2} could co-immunoprecipitate with A3H-HA, suggesting that a lack of Vif-induced degradation did not always correlate with a lack of an interaction. To quantify this interaction, we used rotational anisotropy. First, we determined the oligomerization state of A3H by size exclusion chromatography. In contrast to other single Z-domain A3 enzymes [205, 206], we found that A3H existed as polydisperse forms in solution (Fig. 5.3A). Based on apparent molecular masses, A3H can exist as a monomer (apparent molecular mass of 22 kDa), dimer (apparent molecular mass of 44 kDa), and trimer/tetramer (apparent molecular mass of 67 kDa). The dimeric A3H was the most predominant form (Fig. 5.3A, Fraction 28). Similar to A3G, we found that the rotational anisotropy of fluorescently labeled A3H decreased upon the addition of the Vif heterotetramer (Fig. 5.3B), suggesting that Vif disrupted A3H oligomerization. We found that in the presence of the Vif_{HXB2} heterotetramer, the A3H peak shifted from Fraction 28 to Fractions 22-24 (Fig. 5.3C). The A3H in Fractions 22, 23, and 24 was approximately equal suggesting that Vif_{HXB2} interacted with A3H monomers (Fraction 24), A3H dimers (Fractions 23 and 24) and A3H tetramers (Fraction 22). Thus, the observed decrease in rotational anisotropy (Fig. 5.3C) may be a combination of a change in A3H oligomerization and structural conformation, both of which can cause the molecules to rotate faster [370]. Notably, the apparent K_d of Vif_{HXB2} for A3H was ~60-fold lower than that of Vif_{IIIB} for A3H (Fig. 5.3B, Vif_{HXB2}, apparent K_d 7 ± 2; Vif_{IIIB}, apparent K_d 442 ± 89). Binka *et al.* [240] showed amino acid at position 39 and 48 in Vif are important for interacting with A3H, as mutations F39S and H48N disrupt the interaction of Vif_{LAI} with A3H. Since both Vif_{HXB2} and Vif_{IIIB} have ^{39}F, but lacks a histidine at position 48 and instead encodes an asparagine at position 48 (Fig. 5.3D), the binding data essentially measure the contribution of Vif amino acid 48 in the binding strength of the A3H-Vif interaction [276]. A3H exemplified that although Vif_{HXB2} bound A3H with higher affinity than Vif_{IIIB}, both Vifs bound A3H in the low nanomolar range, but this was not sufficient for complete Vif-mediated degradation of A3H. Therefore, the interaction strength appeared not to be the only determinant.

5.3.3 Vif_{HXB2} can induce degradation of A3G D128K. To further test if the difference in Vif-mediated degradation efficiency is influenced by the A3-Vif interface, we used the A3G D128K

A A3H

B Vif$_{IIIB}$, K$_d$ 442 ± 89 nM ○
Vif$_{HXB2}$, K$_d$ 7 ± 2 nM ●

C A3H-Vif/CBFβ/EloB/C

A3H dimer

Vif/CBFβ/EloB/C

D

```
              16                    39          48
               |                     |           |
Vif_IIIB   MRInTWKrLVKHHMYiSrKAkdWFYRHHYEStnPkISS
Vif_HXB2   MRIrTWKsLVKHHMYvSgKArgWFYRHHYESphPrISS
```
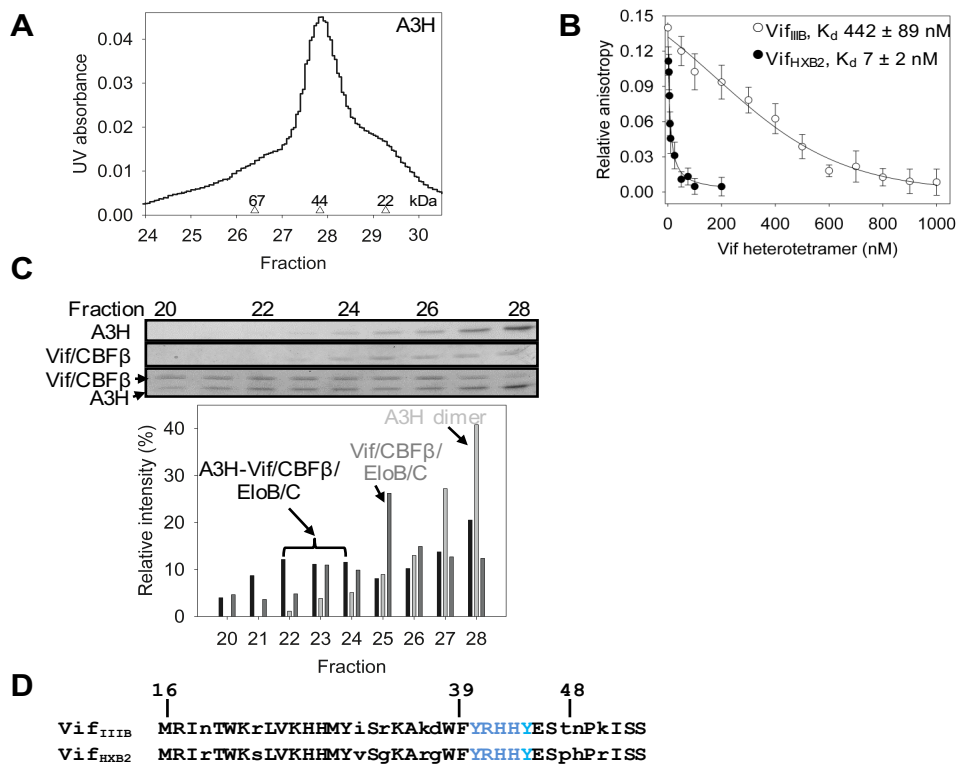
Figure 5.3 A3H interacts with both Vif$_{IIIB}$ and Vif$_{HXB2}$ in vitro. (A) By comparison to a standard curve, size exclusion chromatography was used to determine the oligomerization state of A3H. A3H resolved primarily as a dimer with an apparent molecular mass of 44 kDa with minor trimer/tetramer (67 kDa) and monomer (22 kDa) fractions (open triangles). (B) Steady state rotational anisotropy (fluorescence depolarization) was used to measure the binding strength between A3H and Vif$_{IIIB}$/CBFβ/EloB/C or Vif$_{HXB2}$/CBFβ/EloB/C (labeled as Vif$_{IIIB}$ or Vif$_{HXB2}$). The Vif variant heterotetramer was titrated into a binding reaction with fluorescently labeled A3H to determine an apparent dissociation constant (K$_d$). The Vif$_{HXB2}$ heterotetramer (apparent K$_d$ 7 ± 2 nM) interaction with A3H was of ~60-fold higher affinity than the Vif$_{IIIB}$ heterotetramer (apparent K$_d$ 442 ± 89 nM). Error bars represent the standard deviation of the mean from three independent experiments. (C) Size exclusion chromatography was used to determine the oligomerization state of an A3H-Vif/CBFβ/EloB/C complex. For comparison, A3H alone and the Vif/CBFβ/EloB/C heterotetramer alone were resolved by size exclusion chromatography. Peak fractions were determined by the UV absorbance of the chromatogram and confirmed by resolving samples by SDS-PAGE. The graph represents the quantification of the coomassie stained bands shown in the SDS-PAGE gels. Comparison to molecular weight standards was used to calculate apparent molecular masses of peak fractions. A3H resolved primarily as a dimer (44 kDa, Fraction 28). The Vif/CBFβ/EloB/C heterotetramer resolved as a stoichiometric complex with an apparent molecular mass of 74 kDa in the peak fraction (Fraction 25). The A3H-Vif/CBFβ/EloB/C complex resolved in Fractions 22-24 which gave an apparent molecular mass range of 157-96 kDa. The presence of EloB/C in fractions containing Vif/CBFβ was confirmed by staining with Bio-rad Oriole stain (gel not shown). (D) Alignment of Vif$_{IIIB}$ and Vif$_{HXB2}$ amino acid sequences from amino acids 16 to 53. Residues that differ between the Vif variants are shown in lower case. The [40]YRHHY[44] motif highlighted in blue is the A3G interaction motif. Amino acids 39 and 48 are important for both the ability of Vif to induced degradation of and interact with A3H.
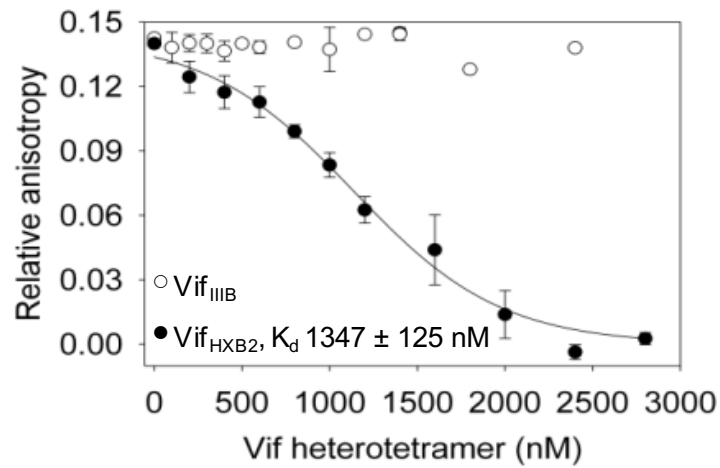
Figure 5.4 D128K interacts with Vif variants. Steady state rotational anisotropy (fluorescence depolarization) was used to measure the binding strength between D128K and $Vif_{IIIB}/CBF\beta/EloB/C$ or $Vif_{HXB2}/CBF\beta/EloB/C$ (labeled as Vif_{IIIB} or Vif_{HXB2}). The Vif variant heterotetramer was titrated into a binding reaction with fluorescently labeled D128K to determine an apparent dissociation constant (K_d). The Vif_{IIIB} heterotetramer was unable to interact with A3G D128K, in contrast to the Vif_{HXB2} heterotetramer (apparent K_d 1347 ± 125 nM).

mutant. D128K is resistant to transiently overexpressed Vif_{IIIB}-mediated degradation [278, 280, 281], but only partially resistant to transiently overexpressed Vif_{HXB2}-mediated degradation. Specifically, Vif_{HXB2} can induce degradation of ~ 40% of the transiently overexpressed D128K [312]. That the degradation was not strong was consistent with results from others [279] and the co-imunoprecipitation data by Baig *et al.* in which neither Vif_{IIIB} nor Vif_{HXB2} could co-immunoprecipitate with the D128K-HA [312]. However, using rotational anisotropy, which is a more sensitive method we found that D128K and Vif_{HXB2} did interact with each other, albeit at low affinity (Fig. 5.4, apparent K_d of 1347 ± 125 nM). The interaction strength of Vif and D128K was ~120-fold less than with A3G, thus necessitating titration of 4-fold more Vif heterotetramer to determine the apparent K_d (compare x-axis, Fig. 5.4 and Fig. 5.2A) and consistent with the suboptimal-induced degradation of D128K by Vif_{HXB2}. There was no interaction detected when A3G D128K was incubated with increasing amounts of Vif_{IIIB} heterotetramer (Fig. 5.4). Altogether, these data demonstrate D128K could be degraded by Vif_{HXB2} despite an interaction in the micromolar range (Fig. 5.4) [312], whereas A3H could not be degraded by Vif_{IIIB} and was degraded by Vif_{HXB2} only 2-fold more than D128K [312], despite a low nanomolar binding affinity (Fig. 5.3B), further indicating that mechanistically the different interfaces of A3G/D128K and Vif_{HXB2} and A3H and Vif_{HXB2} influences degradation efficiency.

5.3.4 Degradation-independent effects of Vif on A3H. For A3H and Vif_{IIIB}, the interaction maintained in the absence of A3H degradation may be a consequence of maintaining a lower affinity interaction to decrease the evolutionary barrier when HIV adapts to rare hosts that express a stable A3H haplotype [100]. Alternatively or in addition to this, Vif_{IIIB} may exert degradation-independent inhibitory effects on A3H through this interaction. Degradation-independent effects of Vif on A3G have been characterized such as inhibition of virion encapsidation, depression of A3G mRNA translation, and inhibition of deamination activity of virion encapsidated A3G molecules that escape Vif-mediated suppression [183, 223, 225, 296, 298, 299, 348, 349]. We did not find that Vif_{IIIB}/CBFβ or Vif_{HXB2}/CBFβ could alter the deamination activity or processive scanning of A3H (Fig. 5.5), as was previously found for A3G (Fig. 3.8) [139]. This may be because the primary interaction residue of A3H (^{121}D) is predicted to be located on helix 4, not on loop 7 as for A3G and is on a face of A3H distal to the amino acids predicted to be involved in A3 enzyme processivity [139, 265].
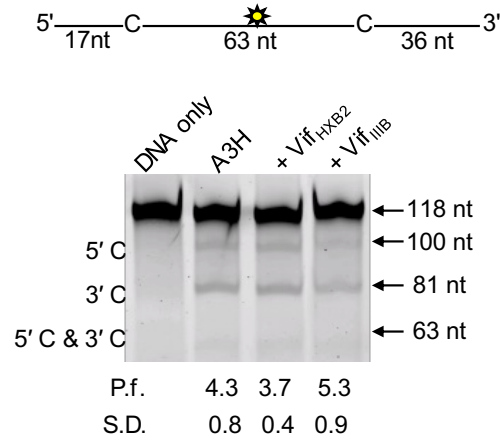
Figure 5.5 Degradation-independent effects of Vif on A3H. Processivity of A3H was tested on a 118 nt single-stranded DNA substrate (100 nM) that contained an internal fluorescein (F)-label (yellow star) and two deamination motifs separated by different distances. The A3H substrates had 5′CTC motifs. The two target cytosines were spaced 63 nt apart. Single deaminations of the 5′C and 3′C were detected as the appearance of labeled 100- and 81- nt fragments, respectively; double deamination of both C residues on the same molecule resulted in a 63 nt labeled fragment (5′C & 3′C). A processivity factor (P.f.) above 1 indicates that the enzyme can catalyze deaminations of both motifs on the substrate in a single enzyme-substrate encounter. Vif_{IIIB} and Vif_{HXB2} cannot decrease the processivity of A3H, in contrast to A3G (see Fig. 3.8) [183]. Due to different specific activities, 50 nM of A3H was used in reactions. The standard deviation of the mean (S.D.) was calculated from three independent experiments.

5.4 Discussion

Due to evidence that different Vif variants have different abilities to degrade specific A3 enzymes [240, 295], we undertook an analysis of various A3-Vif interactions to identify the determinants of an optimal Vif substrate and how the Vif variants differ in their ability to induce A3 degradation. This study is a step towards a better understanding of the A3-Vif relationship at a biochemical level. Based on the enzymes used in our study, we found that there was no clear relationship between the absence of Vif-induced degradation correlating with a lack of an A3-Vif interaction, except for D128K. Rather, the data support a model in which the strength of the interaction between the Vif variant and the A3 enzyme in addition to the A3-Vif binding interface are determinants in Vif-induced degradation efficiency.

Overall, Vif_{HXB2} appears to maintain higher affinity interactions than Vif_{IIIB}. For example, in the absence of the ^{128}D residue, Vif_{HXB2} could maintain an interaction with D128K where Vif_{IIIB} could not (Fig. 5.4). Vif_{HXB2} could interact with A3G and A3H with a 4-fold and 60-fold higher affinity than Vif_{IIIB}, respectively (Fig. 5.2A and Fig. 5.3B). These differences in Vif_{IIIB} and Vif_{HXB2} may relate to the extended A3 binding interface beyond the primary residues important for inducing A3 degradation. The extended A3G binding interface was found to be different for Vif_{IIIB} and Vif_{HXB2} [183]. Experiments examining how Vif variants affected A3G processivity showed that Vif_{HXB2} primarily interacts with loop 7 and Vif_{IIIB} interacts with loop 7 and helix 6 [183]. Therefore, it is conceivable that Vif_{HXB2} could interact with D128K due to stronger interactions with the $^{129}PD^{130}$ residues on loop 7 than Vif_{IIIB}, which maintained the ability of Vif_{HXB2} to induce degradation. The Vif from SIV of Sooty Mangabey Monkeys is able to interact with A3G primarily through ^{129}P [282]. Vif_{IIIB} could not maintain a measurable interaction with D128K (Fig. 5.4) suggesting that it stabilizes itself on A3G through low affinity interactions with helix 6.

The data with A3H and D128K demonstrated that the A3-Vif interface influences degradation efficiency. D128K degradation occurred despite an apparent K_d for Vif_{HXB2} heterotetramer in the micromolar range (Fig. 5.4 and [312]). In contrast, Vif_{HXB2} degradation of A3H was only 2-fold more than for D128K despite a nanomolar range K_d (Fig. 5.3B and [312]). Notably, previously both Vif variants were found to not interact with D128K in a similar *in vitro* assay that used GST-Vif (Table 3.4) [183]. However, our current data with the Vif heterotetramer (Vif/CBFβ/EloB/C) suggests that in the previous study the GST tag

nonspecifically disrupted the Vif_{HXB2} and D128K interaction or further stabilization of Vif with CBFβ/EloB/C was required for the interaction with D128K. However, the GST-Vif produced similar results when interacting with A3G as those obtained here [183], presumably due to the tighter interaction of Vif with wild type A3G (Fig. 5.2A).

5.5 Material and Methods

5.5.1 Protein expression and purification. The GST-A3H was produced using recombinant baculovirus infected *Sf*9 cells. The recombinant baculovirus was constructed using the pAcG2T transfer vector (BD Biosciences) as described in Chelico *et al* [132]. The A3H Haplotype II template used for PCR and subcloning was identical to that used for the pcDNA3.1 vector. The plasmid construction of GST-A3G and GST-D128K has been previously described [132, 139]. The GST-A3H, GST-A3G, and GST-A3G D128K were expressed and purified as described previously to obtain protein that was cleaved from the GST tag and 95% pure (Fig. 5.1) [140].

The Vif heterotetramer (Vif/CBFβ/EloB/C) was obtained by coexpression of all components in *Escherichia coli*. The Vif heterodimer (Vif/CBFβ) was obtained by coexpression of Vif and CBFβ in *E. coli*. The Elongin B (residues 1-187) and Elongin C (residues 17-112) cloned in the pACYC-Duet plasmid were a gift from Dr. Alex Bullock (University of Oxford). The CBFβ isoform 2 (residues 1-182, NCBI Accession NP_001746) was cloned into MCS1 of pET-Duet at BamHI and HindIII sites. Vif_{HXB2} and $Vif_{IIIB,}$ carried in a pAcG2T baculovirus transfer vector as described previously [139] *were* subcloned into the BglII and XhoI sites of the MCS2 of pET-DUET (Novagen). The codon-optimized sequences of Vif_{HXB2} and Vif_{IIIB} were also cloned into the NcoI and HindIII sites of the MCS1 of pRSF-Duet and pCDF-Duet, respectively, to supplement Vif expression. Protein expression in *E. coli* BL21 cells (DE3) cells was induced with 0.2 mM isopropyl-D-thiogalactopyranoside (IPTG) for 16 h at 16 °C. Cells were lysed by sonication in ice-cold lysis buffer (20 mM Tris-Cl, pH 8.0, 300 mM NaCl, RNaseA (80 μg/mL, Qiagen), EDTA-free protease inhibitor (Roche) and lysozyme (1mg/mL, EMD chemicals). Clarified supernatants were transferred to TALON Metal affinity Resin (Clontech Laboratories) and incubated for 45min at 4 °C. The affinity resin was then extensively washed with buffer containing increasing amounts of imidazole (20 mM Tris-Cl, pH 8.0, 300 mM NaCl and 5-10 mM Imidazole). Affinity resin bound proteins were eluted with elution buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 200 mM Imidazole, and 10% glycerol). The eluted proteins were loaded onto a Superdex 200 Increase (10/300) column (GE Healthcare) equilibrated with buffer

containing 20 mM Tris-Cl, pH8.0, 300mM NaCl, 10% glycerol and 1mM DTT. Peak fractions corresponding to the Vif/CBFβ/EloB/C heterotetramer or Vif/CBFβ heterodimer were identified based on comparison to molecular weight standards (Bio-Rad) and confirmed with SDS-PAGE. The Vif heterotetramer was further purified using a DEAE Fast Flow column (GE Healthcare). The proteins were loaded onto the column in low salt buffer (20 mM HEPES, pH 7.0, 75 mM NaCl, 10% glycerol, 1 mM DTT) and eluted using a salt gradient. The Vif heterotetramer eluted from the column at approximately 500 mM NaCl. The complexes were estimated to be 95% pure by SDS-PAGE (Fig. 5.1).

To determine the oligomerization state of A3G, A3H, and Vif_{HXB2} heterotetramer, 100-150 μg of purified enzyme was loaded onto a Superdex 200 Increase (10/300) column (GE Healthcare) equilibrated with buffer containing 20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10% (v/v) glycerol, 1 mM DTT. For determination of the A3G- or A3H- Vif_{HXB2} heterotetramer complex, 100 μg of each A3G or A3H and the Vif_{HXB2} heterotetramer were premixed and incubated at room temperature for 3 min before loading on the column. The Bio-Rad gel filtration standard set was used to generate a standard curve from which the molecular masses and oligomerization state was calculated.

5.5.2 Steady state fluorescence depolarization. The affinity of the interaction between A3G, A3G D128K or A3H and Vif/CBFβ/EloB/C was determined using steady state fluorescence depolarization (rotational anisotropy). The purified A3G, A3G D128K or A3H was labeled with fluorescein using the Fluorescein-EX Protein Labeling kit (Invitrogen) and used as the binding substrate for Vif/CBFβ/EloB/C. Reactions (50μl) were conducted in buffer (50 mM Tris, pH 7.5, 40 mM KCl, 10 mM $MgCl_2$, 1mM DTT) and contained 50 nM fluorescein labeled A3 enzyme and increasing amounts of Vif/CBFβ/EloB/C (0-2800 nM). A QuantaMaster QM-4 spectrofluorometer (Photo Technology International) with a dual emission channel was used to collect data and calculate anisotropy. Measurements were made at 21°C. Samples were excited with vertically polarized light at 495 nm (6-nm band pass) and vertical and horizontal emissions were measured at 520 nm (6-nm band pass). An apparent dissociation constant (K_d) was obtained by fitting to a hyperbolic decay or sigmoidal curve equation using SigmaPlot 11.2 software.

5.5.3 In vitro deamination assay. The 118 nt ssDNA substrates were obtained from Tri-Link Biotechnologies and are listed in Table 5. Substrates containing two 5′CTC motifs (A3H) were used at 100 nM and a fluorescein labeled thymidine was placed in between the motifs.

112

Table 5 Primers and DNA substrates.

Name	Sequence
Vif$_{IIIB}$ Forward pETDuet MCS2	TTTTAGATCTATGGAGAACCGGTGGCAGGTG
Vif$_{IIIB}$ Reverse pETDuet MCS2	TTTTCTCGAGCTAGTGTCCATTCATTGTATGGCTC
Vif$_{HXB2}$ Forward pETDuet MCS2	TTTTAGATCTATGGAAAACAGATGGCAGGT
Vif$_{HXB2}$ Reverse pETDuet MCS2	TTTTCTCGAGCTAGTGTCCATTCATTGTGTG
CBF-β Forward pETDuet MCS1	TTTTGGATCCATGCCGCGCGTCGTGCCCGACCAGA
CBF-β Reverse pETDuet MCS1	TTTTAAGCTTCTAGGGTCTTGTTGTCTTCTTGCCAGTTACTGCCAG
A3G-D128K SDM Forward	CC CGC CTC TAC TAC TTC TGG AAG CCA GAT TAC CAG GAG GCG CTT C
A3G-D128K SDM Reverse	GA AGC GCC TCC TGG TAA TCT GGC TTC CAG AAG TAG TAG AGG CGG G
Vif$_{HXB2}$ Forward pRSFDuet MCS1	TTTTCCATGGAAAACCGCTGGCAAGTTATGAT
Vif$_{HXB2}$ Reverse pRSFDuet MCS1	TTTTAAGCTTTCAATGGCCGTTCATCGTAT
Vif$_{IIIB}$ Forward pCDFDuet MCS1	TTTTCCATGGAAAACCGCTGGCAAGTTATGATT
Vif$_{IIIB}$ Reverse pCDFDuet MCS1	TTTTAAGCTTCCAATGGCCGTTCATCGTATG
A3H 118 nt substrate	GAA TAT AGT TTT TAG CTC AAA GTA AGT GAA GAT AAT [Fam-dT] TAG AGA GTT GTA ATG TGA TAT ATG TGT ATG AAA GAT ATA AGA CTC AAA GTG AAA AGT TGT TAA TGT GTG TAG ATA TGT TAA
A3G 118 nt substrate	GAA TAT ATG TTG AGA CCC AAA GTA ATG AGA GAT TGA [dT-Fam] TAG ATG AGT GTA ATG TGA TAT ATG TGT ATG AAA GAT ATA AGA CCC AAA GAG TAA AGT TGT TAA TGT GTG TAG ATA TGT TAA

Abbreviations are: SDM, site directed mutagenesis; Fam, fluorescein.

For A3H, 50nM of enzyme was incubated with 1750 nM Vif (IIIB or HXB2) during the reaction and the molar ratio of Vif:A3 enzyme was maintained at 35:1. The reaction buffer used was identical to that used for steady state fluorescence depolarization. All reactions were carried out under single hit conditions (~10% substrate consumed) where a single A3 could interact with a particular ssDNA substrate once at most [362]. Reactions were incubated at 37 °C for 1.5 to 30 min. Deaminations were detected by resolving DNA that had been treated with Uracil-DNA Glycosylase and heated under alkaline conditions on a 10% (v/v) denaturing polyacrylamide gel. Gel pictures were obtained using a Typhoon Trio (GE Healthcare) multipurpose scanner, and analysis of integrated gel band intensities was performed using ImageQuant software (GE Healthcare). Under these conditions, a processivity factor can be determined by comparing the total number of deaminations occurring at two sites on the same DNA substrate with a calculated theoretical value of the expected deaminations that would occur at those two sites if the events were uncorrelated, i.e. not processive [132].

6.0 Biochemical characterization of the single-domain A3H enzyme provides insights for oligomerization models within A3 family members

Although A3G has taken center stage among all seven A3 proteins for a long time due to its potent HIV restriction ability, through a few seminal studies the importance of A3H during HIV infection was eventually recognized [45, 100, 232]. A3H is the most polymorphic amongst all human A3 proteins with at least 7 haplotypes (hap I-VII) being reported (Table 7.1). The role each amino acid polymorphism played in the cellular functions of A3H has been examined over the years. Specifically, polymorphic changes at amino acid position 15 and 105 have been reported to be the determinants for the cellular stability of the protein [45, 371]. Amino acid polymorphism of K/D at position 121 determines the sensitivity of A3H to Vif [195, 197, 290]. In this biochemical characterization, I investigated the function of A3H amino acid 178 polymorphism as the polymorphism at this position had no prior ascribed functions [197].

The three stable haplotype (hap II, V and VII) all have N15, R18, R105 and vary at the 121 and 178 positions [45, 198] (Table 1), where A3H hap II, V and VII have a DD, DE and KE sequence at positions 121 and 178, respectively [100, 213, 232]. During the process of biochemically characterizing the determinants required for efficient Vif-mediated degradation of A3G and A3H enzymes (Chapter 5), we have examined the oligomerization states for all three stable haplotypes (hapII, V and VII) and unexpectedly found that the three stable haplotypes demonstrate different oligomeric states in solution. Specifically, A3H hap II and hap V exhibited a similar oligomerization profile as they both primarily form dimers in solution (Fig. 7.5), suggesting amino acid D/E at position 178 on predicted helix 6 was not responsible for mediating dimerization; however, A3H hap VII was able to form an equilibrium of tetramers/pentamers and higher-order oligomers in solution (Fig. 6.0), suggesting amino acid D/K at position 121 near the predicted loop 7 might be involved in larger molecular weight complex formation. Altogether, these results indicate that unlike double-domain A3G which dimerizes through the NTD loop 7 [140] or single-domain Z2-type A3C which dimerizes through helix 6 [205, 216], A3H may dimerize in solution using an interface that is completely distinct from the other A3 family members. Driven by this hypothesis, in Chapter 7 we set out to understand the biochemical basis of A3H oligomerization and to explore how its oligomeric state transitions upon substrate binding.
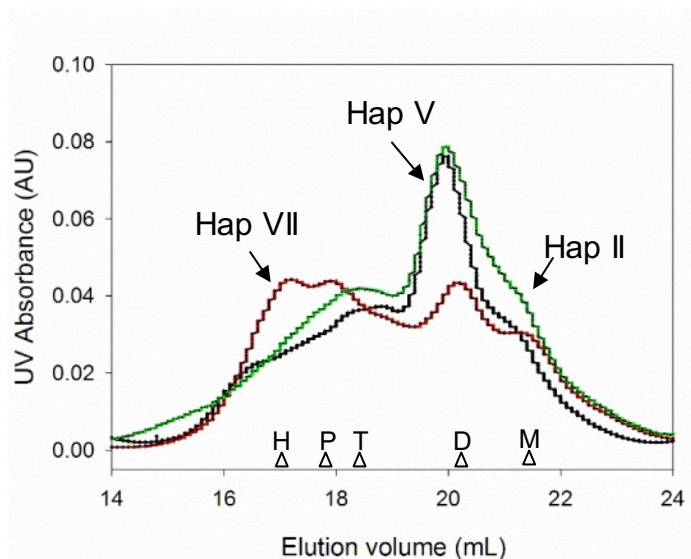
Figure 6.0 Size exclusion chromatography profiles of stable A3H haplotypes. A3H haplotype II (hap II), haplotype VII (hap VII) and haplotype V (hap V) obtained from a 25-mL G200 Superdex Increase column. Hap VII exhibits a different oligomerization profile as compared to hapII and hap V, with dimer (D, 38.2kDa), pentamer (P, 100kDa) and higher order oligomer (H, 163kDa) forming an equilibrium in solution. A3H hap II and hap V demonstrate a similar oligomeric state, as they both mainly form dimers (42 kDa) in the solution, with monomers (22 kDa) and tetramers (94kDa) also been observed. The size exclusion chromatography profile of A3H hap VII, hap V and hap II (300 µg) was used to calculate the oligomerization state of the enzymes from a standard calibration curve (Fig. 7.5, A and B).

116

The experiments in Chapter 7 were carried out by Y.F., with the following exceptions: experiments in Figure 3 and Table 4 were performed by R.P.L. and A.A., experiments for A3H haplotype 5 data in Figure 1 were performed by T.T.B., and experiments for A3H haplotype II and A3G binding data in Table 2 were performed by M.B.A.. Y.F., R.P.L., and L.C. conceived and designed the experiments. Y.F., R.P.L., A.A., M.B.A., T.T.B., and L.C. analyzed the data. Y.F. and L.C. wrote the paper. Y.F., R.P.L., A.A., M.B.A., T.T.B., and L.C. revised the article.

7.0 Natural polymorphisms and oligomerization of human A3H contribute to single-stranded DNA scanning ability

Yuqing Feng[1], Robin P. Love[1], Anjuman Ara[1], Tayyba T. Baig[1], Madison B. Adolph[1], Linda Chelico[1]

[1]Department of Microbiology & Immunology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada, S7N 5E5

Keywords: Deoxycytidine deaminase, processivity, DNA-protein interaction, restriction factor

Running title: A3H polymorphism and oligomerization impact processivity

7.1 Abstract

APOBEC3H is a deoxycytidine deaminase that can restrict the replication of HIV-1 in the absence of the viral protein Vif that induces APOBEC3H degradation in cells. APOBEC3H exists in humans as seven haplotypes (I-VII) with different cellular stabilities. Of the three stable APOBEC3H haplotypes (II, V, and VII), haplotypes II and V occur most frequently in the population. Despite APOBEC3H being a *bona fide* restriction factor, there has been no comparative biochemical characterization of APOBEC3H haplotypes. We characterized the single-stranded (ss)DNA scanning mechanisms that haplotypes II and V use to search their ssDNA substrate for cytosine containing deamination motifs. APOBEC3H haplotype II was able to processively deaminate multiple cytosines in a single enzyme-substrate encounter by using sliding, jumping, and intersegmental transfer movements. In contrast, APOBEC3H haplotype V exhibited diminished sliding and intersegmental transfer abilities, but was able to jump along ssDNA. Due to an Asp or Glu at amino acid 178 differentiating these APOBEC3H haplotypes, the data indicated that this amino acid on helix 6 contributes to processivity. The diminished processivity of APOBEC3H haplotype V did not result in a reduced efficiency to restrict HIV-1 replication in single-cycle infectivity assays, suggesting a redundancy in the contributions of jumping and intersegmental transfer to mutagenic efficiency. Optimal processivity on ssDNA also required dimerization of APOBEC3H through the β2 strands. The findings support a model in which jumping can compensate for deficiencies in intersegmental transfer and suggest that APOBEC3H haplotypes II and V induce HIV-1 mutagenesis efficiently, but by different mechanisms.

7.2 Introduction

The human APOBEC3 (A3) family contains seven deoxycytidine deaminases that act as cellular restriction factors for a number of retrotransposons, retroviruses, and DNA viruses [66, 319]. Excluding A3 polymorphisms, only four of the human A3 enzymes, A3D, A3F, A3G, and A3H are able to restrict the replication of HIV by becoming encapsidated into budding virions and deaminating cytosines to uracils on the nascent (-)DNA synthesized by HIV reverse transcriptase [78, 80](2017 NAR ref). The uracil lesions result in transition mutations upon using the (-) DNA as a template to synthesize the (+) DNA [22-24]. This results in decreases in HIV infectivity that are dependent on the number of mutations and the effect of the mutations on the viral proteins. Each deaminase has a preferred sequence context for deamination, but the

119

minimal dinucleotides for recognition are 5'CC for A3G and 5'TC for A3D, A3F, and A3H [49-52, 91, 195].

The number of mutations that an A3 can introduce into the HIV proviral genome is mediated by the level of encapsidation into the viral capsid and the biochemical properties of the A3 [21, 66]. The primary factor mediating A3 restriction ability is the HIV viral infectivity factor (Vif) protein [25]. Vif acts as the substrate receptor of a Cullin RING ligase-5 (CRL5) E3 ligase complex to induce polyubiquitination and degradation of A3s relevant to HIV restriction [24, 63, 64, 70, 223, 225, 246]. This greatly decreases the encapsidation levels of A3s into HIV virions, but does not completely block A3 virion encapsidation since G to A mutations in a sequence context indicating the action of A3 enzymes are present in HIV genomes recovered from infected individuals [115, 119]. In the absence of Vif and despite equal encapsidation of some A3s into virions, there are differences in their efficiency of HIV restriction, e.g., A3G and A3F [133, 176]. These data demonstrate that there are biochemical differences in the enzymes that influence their activities as restriction factors [133]. We have determined that the mechanisms that A3G and A3F use to scan the single-stranded (ss) DNA substrate in search of the deamination motifs affects the number of induced mutations [133, 139]. Synthesis of the proviral double-stranded (ds) DNA is a dynamic process concurrently involving (-) DNA synthesis, RNA degradation, and synthesis of (+) DNA that blocks A3 access to its single-stranded (-) DNA substrate [4, 91]. As a result, the enzymes require an efficient search process to locate the deamination motifs [66]. A3 enzymes studied thus far have been found to scan ssDNA by facilitated diffusion [93, 132, 133, 206]. Facilitated diffusion can involve 1-dimensional (1-D) sliding movements or 3-D jumping/hopping or intersegmental transfer movements on ssDNA [145, 146]. The most efficient search on ssDNA would involve a combination of 1-D and 3-D scanning movements [66, 139, 372]. The 1-D sliding motions allow for an in-depth local search, i.e., < 20 nt, of the substrate for the deamination motif. The 3-D jumping/hopping or intersegmental transfer movements mediate larger translocations on the DNA substrate [148]. Each type of movement provides a superficial scanning mechanism by itself, but together the enzyme is able to transition between an in-depth search (sliding) for the catalytic sequence to random relocation (jumping/hopping or intersegmental transfer) on the DNA substrate to restart the search process [133, 139]. Jumping/hopping are differentiated from intersegmental transfer by the type of association the enzyme has with the DNA. Jumping/hopping involves microscopic dissociation and re-

120

association of the enzyme with the same DNA without diffusion into the bulk solution. Intersegment transfer involves an enzyme with two DNA binding domains engaging in a doubly-bound state with the DNA substrate before moving to one of the bound sites [146]. Facilitated diffusion movements enable enzyme processivity where an enzyme can catalyze multiple reactions before dissociating from the substrate and diffusing back into the bulk solution [66].

Of the A3 enzymes most effective against HIV, A3H is of special interest since it occurs with the highest number of haplotypes in the A3 family [45, 195, 198]. A3H exists as at least seven haplotypes in humans, with three of the haplotypes being stable forms of A3H that can suppress HIV (Table 7.1). The A3H haplotypes that are unstable and inactive against HIV contain two single nucleotide polymorphisms that occur in different combinations with other polymorphisms. Unstable A3H haplotypes (haplotypes I, III, IV, and VI) carry either a G105, Δ N15, or both in combination with other possible polymorphisms at positions 18, 121, and 178 (Table 7.1) [45, 198]. Unstable A3H haplotype I occurs at the highest population frequency (0.308-0.526) [45, 198]. A R105 and N15 are required for a stable A3H in cells [45]. The three stable haplotypes all have a N15, R18, R105 and vary at the 121 and 178 positions (Table 7.1) [45, 198]. A3H haplotype II has been the most widely studied and has a DD sequence at positions 121 and 178 [100, 213, 232]. A3H haplotype V has a DE sequence at positions 121 and 178 [195, 198]. A3H haplotype II and V comprise the majority of the stable A3H alleles (haplotype II, frequency of 0.061-0.265; haplotype V, frequency of 0.054-0.202) [45, 198]. A3H haplotype VII has been identified only from one individual (frequency of 0.009) and has a KE sequence at positions 121 and 178 [198]. The 121 position determines the sensitivity to Vif, where a D121 is sensitive and a K121 is not sensitive to Vif-mediated degradation [195, 197, 290]. There is currently no known function of the polymorphism of D/E at position 178 [197].

Despite the several unique aspects of A3H in comparison to the other A3 enzymes, there has been no in depth biochemical characterization of the most frequently occurring stable A3H haplotypes (II and V). This type of characterization is necessary to fully understand the determinants required for HIV restriction since A3H has several distinctive characteristics in the A3 family. A3s can be divided into two families based on the number of Zinc coordinating domains (Z-domains) [37]. A3B, A3D, A3F and A3G are deaminases with two Z-domains, whereas A3A, A3C and A3H are deaminases with a single Z-domain. In addition, A3s have been classified into three categories based on the phylogenetic analysis of their Z-domains (Z1, Z2

Table 7.1 A3H haplotypes.

Haplotype	Stability	Polymorphisms (amino acids 15, 18, 105, 121, 178)
I	Unstable	NRGKE
II	Stable	NRRDD
III	Unstable	ΔRRDE
IV	Unstable	ΔLRDE
V	Stable	NRRDE
VI	Unstable	ΔLGKE
VII	Stable	NRRKE

Haplotype data summarized from OhAinle *et al*. and Wang *et al* [45, 198]

and Z3) and A3H is the only Z3-type domain A3 enzyme [37]. Part of the reason why A3H is able to restrict HIV is that it is the only wild-type single Z-domain A3 that oligomerizes in solution and on cellular RNA, which is generally prerequisite for virion encapsidation [191, 312]. Importantly, in another single Z-domain A3, A3C, a rare S188I polymorphism that occurs in 10% of people of African descent, enables A3C dimerization and this results in A3C S188I being able to restrict HIV [216]. Thus, the oligomerization of single Z-domain A3 enzymes appear to have a large effect on their ability to contribute to cellular functions.

We undertook a biochemical characterization of A3H haplotypes II and V in order to determine if there are shared biochemical characteristics between single Z-domain and double Z-domain A3 enzymes for restriction of HIV. The data demonstrate that A3H is a processive enzyme. A3H haplotype II uses sliding, jumping and intersegmental transfer to locate deamination motifs on ssDNA. The E178 polymorphism of A3H haplotype V causes diminished sliding and intersegmental transfer, but does not affect jumping. This differences in processivity did not affect HIV restriction ability. Further, we identified that the dimerization interface in A3H is mediated by the pairing of β2 strands from two monomers and tetramerization can occur through the pairing of loop 7 from two dimers, as predicted could occur from a crystal structure of APOBEC2 (A2) with a 40 amino acid N-terminal deletion [373]. The β2-β2 strand dimerization of A3H is required for efficient ssDNA scanning. The results demonstrate that A3H dimerization and D178 is required for sliding and intersegmental transfer movements on ssDNA. The data combine to support a model in which A3H haplotypes II and V both induce HIV mutagenesis efficiently, but by different mechanisms and suggest that jumping can compensate for deficiencies in intersegmental transfer.

7.3 Results
7.3.1 A3H Haplotype II and V processively scan ssDNA. The ssDNA scanning mechanism(s) of A3H haplotypes II and V were determined using an *in vitro* deamination assay that measures the processivity of purified A3H. Processivity is defined as the ability of an enzyme to deaminate more than one cytosine on an ssDNA substrate in a single enzyme-substrate encounter [132]. The processivity was measured using different synthetic ssDNA substrates containing two deamination motifs (5'CT<u>C</u>, underlined C being preferentially deaminated) spaced various
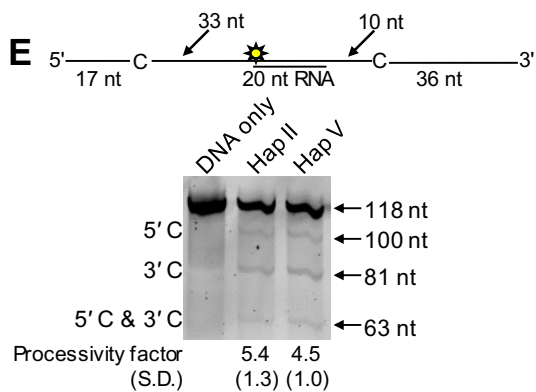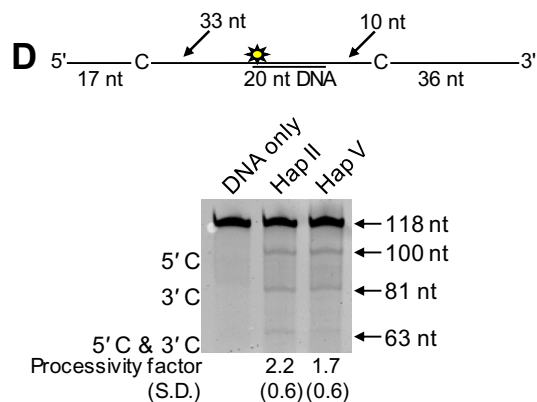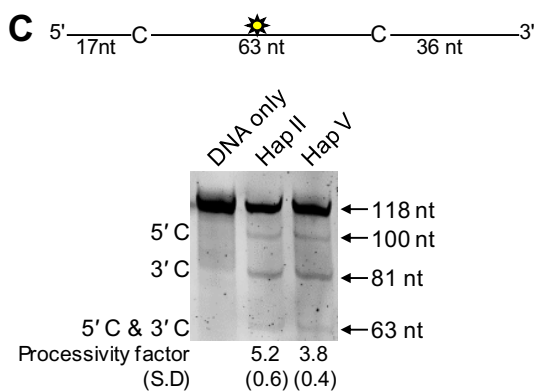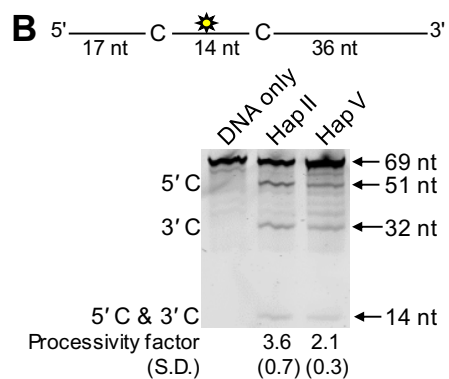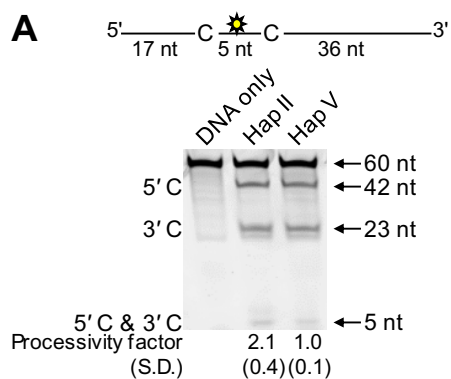
Figure 7.1 Analysis of A3H haplotypes II and V processivity and ssDNA scanning. Processivity of A3H haplotype II (Hap II) and haplotype V (Hap V) was tested on ssDNA substrates that contained fluorescein labeled deoxythymidine (yellow star) between two 5'CTC deamination motifs separated by different distances. (A) Deamination of a 60 nt ssDNA substrate with deamination motifs spaced 5 nt apart. Single deaminations of the 5'C and 3'C are detected as the appearance of labeled 42 nt and 23 nt fragments, respectively; double deamination of both C residues on the same molecule results in a 5 nt labeled fragment. (B) Deamination of a 69 nt ssDNA substrate with deamination motifs spaced 14 nt apart. Single deaminations of the 5'C and 3'C are detected as the appearance of labeled 51 nt and 32 nt fragments, respectively; double deamination of both C residues on the same molecule results in a 14 nt labeled fragment. (C) Deamination of a 118 nt ssDNA substrate with deaminated cytosines spaced 63 nt apart. Single deaminations of the 5'C and 3'C are detected as the appearance of labeled 100 nt and 81 nt fragments, respectively; double deamination of both C residues on the same molecule results in a 63 nt labeled fragment. (D). Deamination of the same substrate described for (C), but with a 20 nt DNA annealed between the two deamination motifs. (E) Deamination of the same substrate described for (C), but with a 20 nt RNA annealed between the two deamination motifs. (A-E) The A3H: ssDNA ratio was 1:2. The measurements of enzyme processivity (Processivity factor) and the Standard Deviation of the mean (S.D.) are shown below the gels. All values are calculated from at least three independent experiments.

distances apart. This strategy allows the dissection between 1-D sliding and 3-D translocation movements since it has been previously shown that substrates with closely spaced motifs, i.e., < 20 nt, are processively deaminated through 1-D sliding motions [133, 139]. Processive deamination of more distantly spaced motifs requires a 3-D translocation mechanism such as jumping or intersegmental transfer [66, 133, 139]. All deamination reactions were performed under "single-hit" conditions (<15% substrate usage) to ensure any given ssDNA substrate was acted upon by at most one enzyme during the reaction. Under these conditions a processivity factor can be calculated to determine the likelihood that an enzyme would undergo processive deaminations in comparison to a nonprocessive deamination (see Material and Methods and [132]).

On ssDNA substrates with deamination motifs separated by 5- and 14- nt, A3H haplotype II had processivity factors of 2.1 and 3.6, respectively (Fig. 7.1, A and B). Thus, depending on the distance between the deamination motifs, A3H haplotype II is ~2 to 4 times more likely to catalyze a processive deamination of the two deamination motifs than a single nonprocessive deamination of one of the deamination motifs. Due to the short distance between the deamination motifs, the data show that A3H haplotype II can processively scan the ssDNA by sliding. The amino acid sequence of A3H haplotype II and A3H haplotype V differ only at the amino acid at position 178 where A3H haplotype II has an Asp and A3H haplotype V has a Glu [195, 198]. Because the amino acid at position 178 is located on predicted helix 6, which mediates sliding ability in A3G and A3F [133, 139], we tested A3H haplotype V on ssDNA substrates with closely spaced deamination motifs to determine if the sliding ability was compromised. On the substrate with deamination motifs separated by 5 nt, the processivity factor was 1 (Fig. 7.1A). Since the processivity factor is a ratio, the processivity factor of 1 means that A3H haplotype V is essentially not processive on this substrate (see Material and Methods), in contrast to A3H haplotype II (Fig. 7.1A). On the substrate with cytosines separated by 14 nt, A3H haplotype V is processive but has a processivity factor ~2-fold less than A3H haplotype II (compare processivity factors of 3.6 and 2.1 (Fig. 7.1B)). The data indicated that in A3H, predicted helix 6 mediates sliding and that the D/E178 polymorphism between A3H haplotype II and haplotype V caused a difference in the ssDNA sliding ability.

We also tested if other processive mechanisms were affected by the polymorphism at amino acid position 178. On the ssDNA substrate with cytosines separated by 63 nt, A3H haplotype V had a processivity factor of 3.8, which is ~1.5-fold less processive as compared to A3H haplotype II on the same substrate (compare processivity factors of 5.2 and 3.8 (Fig. 7.1C)). That both A3H haplotypes were processive on a substrate with distantly spaced deamination motifs suggests that A3H uses a 3-D translocation mechanism to scan ssDNA. To confirm this, we annealed a 20 nt complementary DNA between the two deamination motifs (Fig. 7.1D, *schematic*). Since A3H haplotype II does not bind to this dsDNA region (Table 7.2, no binding detected) it serves as a block to inhibit the sliding motion of the enzyme. A3H haplotype II can still processively deaminate the cytosines on this substrate, (Fig. 7.1D, processivity factor of 2.2), which confirms that A3H haplotype II uses a 3-D translocation mechanism to overcome obstacles while scanning ssDNA. The approximately 2.5-fold decrease in processivity in comparison to the fully ssDNA (compare processivity factors, Fig. 7.1C and 7.1D) is due to attempts to slide over the dsDNA region that instead result in dissociation of the enzyme into the bulk solution. A3H haplotype V bound to dsDNA with an affinity 10-fold less than ssDNA, suggesting that it would also not be able to slide over the dsDNA (Table 7.2, apparent K_d of 7.3 μM, dsDNA and 0.73 μM, ssDNA). Similar to A3H haplotype II, the dsDNA region caused a decrease in the processivity of A3H haplotype V, but the enzyme was still processive (Fig. 7.1D, processivity factor of 1.7). Interestingly, we found that A3H haplotype II and V are able to bind RNA/DNA hybrids with an apparent dissociation constant that is equivalent to that of ssDNA (Table 7.2, A3H haplotype II apparent K_d values of 0.65 μM (RNA/DNA) and 0.72 μM (ssDNA); A3H haplotype V apparent K_d values of 0.82 μM (RNA/DNA) and 0.73 μM (ssDNA)). This is in contrast to A3G that binds both dsDNA and RNA/DNA hybrids at least 3-fold less than ssDNA (Table 7.2, apparent K_d values of 5.70 μM (dsDNA); 0.99 μM (RNA/DNA); and 0.31 μM (ssDNA)). These data predict that A3H should be able to slide over RNA/DNA hybrids. To test this prediction, we annealed a 20-nt complementary RNA between the two deamination motifs (Fig. 7.1E, *schematic*). The enzyme processivity was not significantly different on this substrate in comparison to its analogous ssDNA for both A3H haplotypes II and V (Fig. 7.1, C and E, processivity factors). This suggests that during HIV reverse transcription A3H haplotypes II and V could slide on RNA/DNA hybrid regions.

127

Table 7.2 Comparison of apparent dissociation constants (K_d) from DNA of A3H haplotypes II, V, and A3G.

Enzyme	$K_d, \mu M \pm$ S.D.		
	ssDNA	dsDNA	RNA/DNA
A3H haplotype II	0.72 ± 0.14	No binding detected	0.65 ± 0.09
A3H haplotype V	0.73 ± 0.07	7.30 ± 2.30	0.82 ± 0.29
A3G	0.31 ± 0.04	5.70 ± 1.10	0.99 ± 0.11

To determine the 3-D translocation mechanism of A3H haplotypes II and V, we tested whether the enzyme could move by intersegmental transfer. This mechanism requires that an enzyme with two DNA binding domains transfer to distal sites through a "doubly-bound" intermediate state [146]. A3H haplotype II has been shown to form dimers in solution [66], indicating that it could use intersegmental transfer. To observe intersegmental transfer we increased the enzyme and substrate concentrations in the reaction but kept their ratios constant (Fig. 7.2). By crowding the reaction in this manner the enzyme will become more likely to translocate to a different substrate than to translocate within a single ssDNA as the concentration of the reaction components increases [374]. Thus, if intersegmental transfer is occurring, this would lead to an apparent decrease in the processivity factor. Intersegmental transfer should also result in an increase in the apparent reaction rate since increasing the cycling of the enzyme also increases the rate of searching for deamination motifs [374]. The processivity factor of A3H haplotype II decreased ~2-fold with increasing reaction components (Fig. 7.2A, processivity factor decreases from 5.2 to 3.0) and this was accompanied by a ~2-fold increase in the rate of deamination (Fig. 7.2A, rate increases from 0.35 to 0.6%/min), demonstrating that A3H haplotype II can use intersegmental transfer to scan ssDNA. However, at the highest level of reaction components the processivity plateaued and did not decrease to 1 which would otherwise indicate a complete loss of apparent processivity due to the reaction conditions (Fig.7.2A, Processivity factor of 3.0). This indicates that A3H haplotype II may also use jumping to translocate on ssDNA. Processive ssDNA scanning mediated by jumping is insensitive to the crowding of the reaction components since during this movement the enzyme remains within the negatively charged region of the ssDNA and out of the bulk solution [66]. All together the results indicate that A3H haplotype II uses sliding and both intersegmental transfer and jumping as a 3-D translocation mechanism to scan ssDNA for its preferred deamination motif. We next examined if A3H haplotype V was capable of intersegmental transfer. A3H haplotype V did have a 1.5-fold decrease in apparent processivity, similar to A3H haplotype II, but there was no increase in the reaction rate observed (Figs. 7.2, compare A and B). Thus, the data indicated that although A3H haplotype V could transfer to other ssDNA substrates, there was no corresponding increase in cycling between the ssDNAs, suggesting that the Glu-178 polymorphism also diminished intersegmental transfer ability. The processivity factor from the intersegmental

129
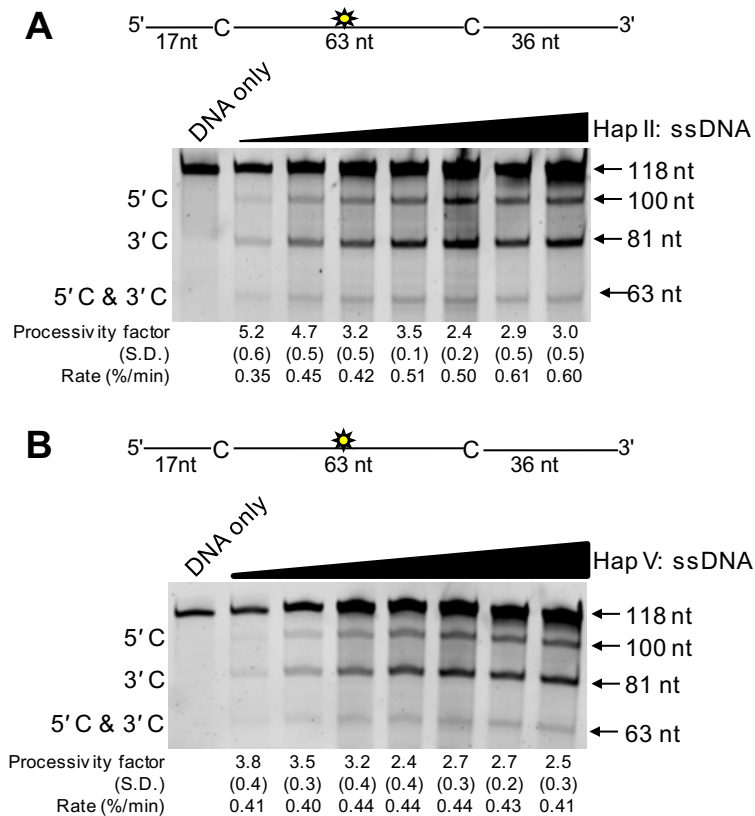
Figure 7.2 A3H can scan ssDNA using intersegmental transfer. Processivity of A3H haplotype II (Hap II) and haplotype V (Hap V) was tested on ssDNA substrates that contained fluorescein labeled deoxythymidine (yellow star) between two 5'CTC deamination motifs separated by different distances. (A-B) Deamination of a 118 nt ssDNA substrate with deaminated cytosines spaced 63 nt apart. Single deaminations of the 5'C and 3'C are detected as the appearance of labeled 100 nt and 81 nt fragments, respectively; double deamination of both C residues on the same molecule results in a 63 nt labeled fragment. For A3H Hap II (A) and A3H Hap V (B) the A3H: ssDNA ratio of 1:2 the reaction components were increased (50 nM: 100 nM; 100 nM: 200 nM; 150 nM: 300 nM; 200 nM: 400 nM; 250 nM: 500 nM; 300nM: 600 nM; 350 nM: 700 nM). These reaction conditions enabled the investigation of whether A3H Hap II and Hap V could translocate on ssDNA by intersegment transfer. The measurements of enzyme processivity (Processivity factor), Standard Deviation of the mean (S.D.), and rate of deamination (percent (%)/min) for each individual reaction condition are shown below the gel. All values are calculated from at least three independent experiments.

130

transfer assay at the highest concentration of reaction components where the decrease in processivity plateaued was greater than 1, suggested that A3H haplotype V was able to processively jump on ssDNA (Fig. 7.2B). Of note, the identity of the amino acid at position 178 also affected the binding of dsDNA and RNA. Although A3H haplotypes II and V did not differ in their ssDNA binding ability as determined by their apparent dissociation constants (K_d) (Table 7.2, haplotype II K_d of 0.72 μM; haplotype V K_d of 0.73 μM), A3H haplotype V was able to bind double-stranded (ds) DNA whereas A3H haplotype II was not able (Table 7.2, haplotype II, no binding detected; haplotype V, K_d of 7.3 μM). There were also haplotype differences in RNA binding. The apparent K_d of A3H haplotype II binding to HIV 5'UTR RNA was 21-fold lower than A3H haplotype V (Table 7.3, haplotype II K_d of 0.22μM; haplotype V K_d of 4.60 μM). This difference in RNA binding ability was consistent with binding data for 7SL Alu RNA where A3H haplotype II bound with an apparent K_d 9-fold lower than A3H haplotype V (Table 7.3, haplotype II K_d of 0.33 μM; haplotype V, K_d of 3.00 μM). Interaction with either HIV or host RNAs such as 7SL, in a nucleocapsid dependent manner, have been reported to be required for encapsidation of A3G and A3H [83, 86, 89, 130, 136, 172, 198, 375, 376]. Similar to A3G [140] and in agreement with previous data [191], A3H bound RNA cooperatively demonstrating the ability to oligomerize on RNA (Table 7.3, Hill coefficients).

7.3.2 A3H Haplotype II and V demonstrate similar HIV-1ΔVif restriction efficiencies. Excluding A3 polymorphisms, human A3H is the only single Z-domain A3 enzyme that encapsidates into budding HIV virions and restricts the replication of HIV in the subsequent target cell [78]. It has been established that A3H haplotype II and haplotype V can similarly restrict the replication of HIV in cell culture or infected individuals [45, 100, 195, 199, 232, 289, 377]. In agreement with others, our data indicated that in a single-cycle of replication, A3H haplotype II and V suppress HIV similarly over a range of A3H expression levels [198, 211, 212, 232] (Fig. 7.3, A and B). Despite the 10 to 20-fold differences in RNA binding affinities between A3H haplotypes II and V (Table 7.3), their virion encapsidation was comparable (Fig. 7.3B). It has been shown that A3G binds cellular or viral RNA indiscriminately to ensure encapsidation into virions in a nucleocapsid specific manner[378] [379] and our data suggest that the same encapsidation mechanism applies to A3H.

It has been previously established for A3G and A3F that an efficient processive DNA

A

B

C

```
HapII_09   TATTAGAAGAAATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAG
HapII_07   TATTAGAAAAAATGAATTTGCCAGAAAAATGGAAACCAAAAATAATAGGGGGAATTGGAG
HapV_01    TATTAAAAAAAATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGAAATTGAAG
HapV_05    TATTAGAAAAAATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGAAG
Reference  TATTAGAAGAAATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATTGGAG
           *****  **  *************** ** *************** ******* ***** **

HapII_09   GTTTTATCAAAGTAAGACAGTATGATCAAATACTCATAGAAATCTGCGGACATAAAGCTA
HapII_07   GTTTTATCAAAGTAAGACAGTATAATCAGATACTCATAGAAATCTGCGGACATAAAGCTA
HapV_01    GTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAAAAATCTGCGAACATAAAGCTA
HapV_05    GTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGCGGACATAAAGCTA
Reference  GTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGCGGACATAAAGCTA
           ***********************  **** ********* ********* ***********
```

Figure 7.3 A3H haplotypes II and V demonstrate similar HIVΔVif restriction efficiencies. (A) HIV ΔVif infectivity was measured by eGFP expression in 293T cells infected with HIVΔVif that was produced in the absence or presence of untagged A3H haplotype II (Hap II) or A3H haplotype V (Hap V). A titration of A3H plasmid was used (25, 50, 100, or 200 ng). Results normalized to the no A3 condition are shown with the Standard Deviation of the mean calculated from at least three independent experiments. (B) The level of HA-tagged A3H Hap II and A3H Hap V expressed in cells and encapsidated into HIVΔVif virions was determined using immunoblotting. The A3 enzymes were detected through the HA tag. The loading control for cell lysates was α-tubulin (α-tub) and for virions was p24. A representative blot from three independent experiments is shown. (C) Proviral DNA sequencing data from a subset of clones (from 100 ng A3H plasmid transfection) showing the sequence context of mutations induced by Hap II or Hap V in the protease region. Asterisks denote homology. Guanine (G) to adenine (A) mutations are shown in boldface.

Table 7.3 Comparison of apparent dissociation constants (K_d) and Hill coefficients from RNA of A3H haplotypes II and V.

Enzyme	K_d, μM $\pm$ S.D. (Hill coefficient)	
	HIV 5′UTR	Alu
A3H haplotype II	0.22 ± 0.04 (2.1)	0.33 ± 0.05 (3.2)
A3H haplotype V	4.60 ± 0.38 (1.8)	3.00 ± 0.60 (2.3)

scanning mechanism involving both 1-D sliding and 3-D translocation is required for inducing high numbers of deoxycytidine deaminations in HIV (-) DNA [133]. Because we observed that A3H haplotypes II and V were encapsidated into virions at similar levels to each other and there were no significant differences in suppression of HIV infectivity (Fig. 7.3, A and B), the data support that the diminished processivity of A3H haplotype V in comparison to A3H haplotype II (Fig. 7.1) is inconsequential to A3H-mediated HIV restriction. This may be because A3H haplotype V did not have a complete loss of sliding ability and still maintained jumping as a 3-D translocation mechanism (Fig. 7.1). The A3H haplotype II data suggest that the 3-D translocation mechanisms are redundant since A3H haplotype II did not restrict HIV replication more than A3H haplotype V over a range of cellular expression levels (Fig. 7.3, A and B). We sequenced proviral DNA from the transfection conditions that used 100 ng of A3H expression plasmid and found no significant differences in the spacing of the A3H haplotype II and V mutations, confirming that the diminished processivity of A3H haplotype V did not affect its ability to induce mutations in HIV proviral DNA (Fig. 7.3C). Processivity can be inferred from the mutational spectra because clustered mutations are indicative of sliding and distantly spaced mutations are indicative of jumping or intersegmental transfer [133, 139, 183]. Because there were similar distantly spaced mutations for both A3H haplotypes the data support that intersegmental transfer and jumping are redundant 3-D processive mechanisms (Fig. 7.3C). Both A3H haplotypes II and V also induced closely spaced mutations, indicative of sliding, in agreement with the processivity data from ssDNA oligonucleotides (Fig. 7.3C and Fig. 7.1 (A and B)). The ~2-fold lower processivity of A3H haplotype V in comparison to A3H haplotype II for closely spaced deamination motifs on ssDNA oligonucleotide substrates (Fig. 7.1, A and B) did not appear to affect its ability to induce closely spaced mutations during proviral DNA synthesis (Fig. 7.3C). At most, there was only a 2-fold difference in the G to A mutation frequency of A3H haplotypes II and V and this was observed only at the 100 ng A3H expression plasmid condition (Fig. 7.3A and Table 7.4).

7.3.3. Optimal sliding and intersegmental transfer requires an Asp-178. The A3H haplotypes II and V differ by only an Asp/Glu at amino acid position 178. These amino acids have differences in side chain rigidity and pK_a that could account for the effect on ssDNA scanning ability.

Table 7.4 Analysis of A3-induced mutagenesis of *prot* DNA from integrated HIV△Vif.

Enzyme	Base pairs sequenced	Total number of G→A mutations	G→A mutation frequency (mutations/kb)
A3H haplotype II	10530	35	3.2
A3H haplotype V	8424	14	1.7

The proviral DNA analyzed was obtained from the transfection condition with 100 ng of A3H expression plasmid.
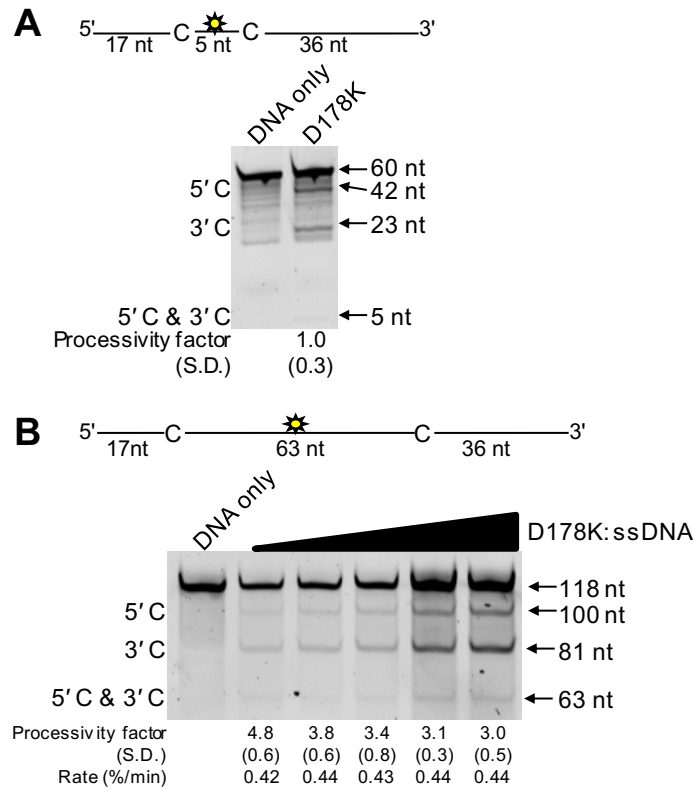
Figure 7.4 A3H D178K is deficient in sliding and intersegmental transfer. Processivity of A3H D178K (D178K) was tested on ssDNA substrates that contained fluorescein labeled deoxythymidine (yellow star) between two 5'CTC deamination motifs separated by different distances. (A) Deamination of a 60 nt ssDNA substrate with deamination motifs spaced 5 nt apart. Single deaminations of the 5'C and 3'C are detected as the appearance of labeled 42 nt and 23 nt fragments, respectively; double deamination of both C residues on the same molecule results in a 5 nt labeled fragment. (B) Deamination of a 118 nt ssDNA substrate with deaminated cytosines spaced 63 nt apart. Single deaminations of the 5'C and 3'C are detected as the appearance of labeled 100 nt and 81 nt fragments, respectively; double deamination of both C residues on the same molecule results in a 63 nt labeled fragment. For D178K the A3H: ssDNA ratio of 1:2 the reaction components were increased (50 nM: 100 nM; 100 nM: 200 nM; 150 nM: 300 nM; 200 nM: 400 nM; 250 nM: 500 nM). These reaction conditions enabled the investigation of whether A3H Hap II and Hap V could translocate on ssDNA by intersegment transfer. The measurements of enzyme processivity (Processivity factor), Standard Deviation of the mean (S.D.), and rate of deamination (percent (%)/min) for each individual reaction condition are shown below the gel. All values are calculated from at least three independent experiments.

136

However, we investigated if a charge reversal or removal mutation could more drastically alter the processivity of A3H. The sliding ability of a D178K mutant was tested using a ssDNA substrate that had deamination motifs separated by 5 nt (Fig. 7.4A). Similar to A3H haplotype V, the A3H D178K mutant was not processive on this substrate, indicating a deficiency in sliding ability (compare Fig. 7.4A and Fig. 7.1A, processivity factors of 1). The D178K mutant was processive on the ssDNA substrate with deamination motifs separated by 63 nt, indicating that it could use a processive 3-D translocation mechanism (Fig. 7.4B). In the assay that tests intersegmental transfer ability, the D178K mutant processivity decreased 1.6-fold with increasing concentration of enzyme and substrate, but there was no increase in the reaction rate (Fig. 7.4B). Thus, the D178K mutant was similar to A3H haplotype V and could transfer to other ssDNA substrates but was unable to cycle efficiently (compare Fig. 7.2B and 7.4B). All together the data indicated that rather than the amino acid charge, the Asp at amino acid position 178 was specifically important for sliding and intersegmental transfer. However, a charged amino acid is important at amino acid position 178 for enzyme stability since a D178A mutant was less stable than A3H haplotype II and had processivity factors of 1.7 or less on all ssDNA substrates tested (data not shown). This suggests that a salt bridge may be formed from the amino acid position at 178 to stabilize the enzyme and influence interactions with ssDNA.

7.3.4 The dimer interface of A3H is mediated by predicted β-β strand interactions. The finding that an amino acid change on A3H predicted helix 6 affected sliding motions was consistent with published data [133, 139], but a role for helix 6 in intersegmental transfer was unexpected. The current understanding of intersegmental transfer is that two DNA binding domains are required [146, 148]. Additional determinants for this ssDNA scanning mechanism have not been previously identified. Yet, there appears to be additional determinants required for A3 enzymes since those with two Z-domains in a single polypeptide, i.e., A3F and A3G, are unable to readily undergo intersegmental transfer under our assay conditions [133]. Based on intersegmental transfer requiring two DNA binding domains [148], the data indicated that since we could not completely abolish intersegmental transfer ability by making mutations at the amino acid at position 178, the A3H was remaining in its oligomeric state and we had identified an amino acid determinant required for intersegmental transfer. To fully characterize the determinants for intersegmental transfer in A3H, we characterized the oligomeric interfaces of A3H haplotype II

A

B

C

β2 strand

Hap II ^{41}-TPTRGYFENK^{50}
A2 ^{83}VQASRGYLEDE^{93}

44 46

D

E

F

GST-Hap II   7.6 mL, 77 kDa
6 mL   8.6 mL

GST-R44A/Y46A   8.3 mL, 37 kDa
6 mL   8.6 mL

G

GST   8.0 mL, 51 kDa
6 mL   8.6 mL
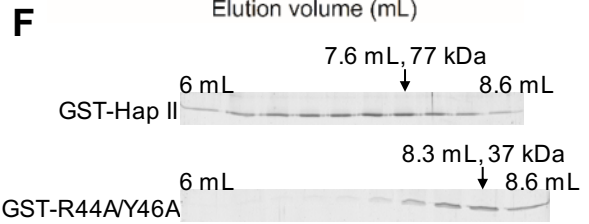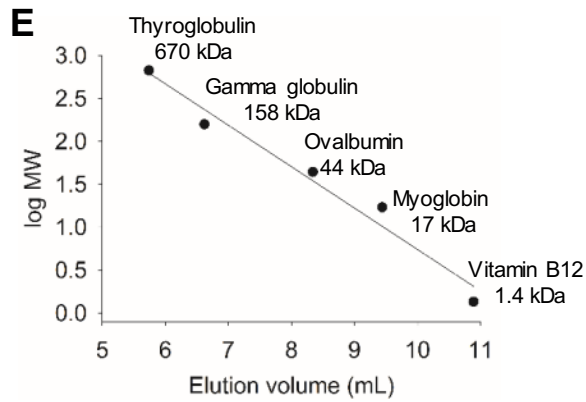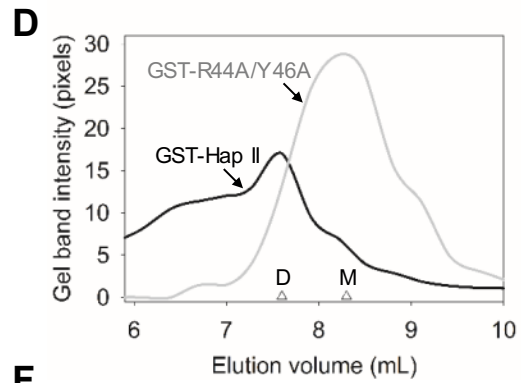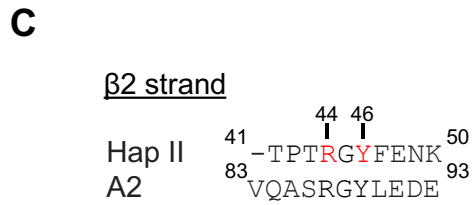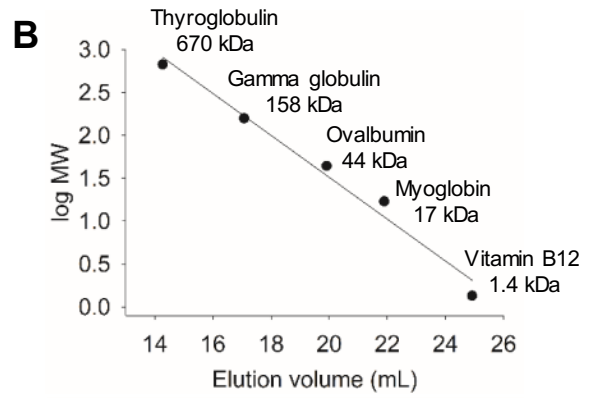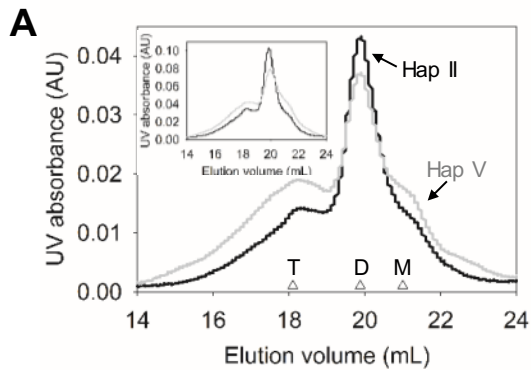
Figure 7.5 β2-β2 strand amino acids mediate A3H dimerization. (A) The size exclusion chromatography profiles of A3H haplotype II (Hap II) and A3H haplotype V (Hap V) obtained from a 25 mL G200 Superdex Increase column were used to calculate the oligomerization states of the enzymes from a (B) standard calibration curve. Analysis demonstrated that both A3H haplotype II (Hap II) and A3H haplotype V (Hap V) were able to form monomers (M), dimers (D) and tetramers (T) in solution. Both 150 μg and 300 μg (inset graph) of enzyme were resolved to investigate whether A3H tetramer formation was concentration-dependent. According the calibration curve the apparent molecular weights of peak fractions for A3H Hap II were 24 kDa (monomer), 44 kDa (dimer), and 94 kDa (tetramer) and for A3H Hap V were 23 kDa (monomer), 44 kDa (dimer), and 102 kDa (tetramer). The same data also appeared in Figure 6.1. (C) Sequence alignment of A3H Hap II and APOBEC2 (A2) β2 strand amino acid sequences. Amino acids that were mutated are indicated in red. (D) The size exclusion chromatography profiles of GST-A3H haplotype II (GST-Hap II) and GST-A3H haplotype II R44A/Y46A (GST-R44A/Y46A) obtained from a 10 mL G200 Superdex column were used to calculate the oligomerization states of the enzymes from a (E) standard calibration curve. When 10 μg of enzyme was loaded onto the size exclusion column, GST-A3H haplotype II formed dimers in solution (apparent molecular weight of 77 kDa in peak fraction). This is in contrast to GST-A3H haplotype II R44A/Y46A, which formed monomers (apparent molecular weight of 37 kDa in peak fraction). (F) The chromatograms from the 10 mL Sephadex 200 column (panel D) were constructed by analyzing the integrated gel band intensities of the protein in each fraction after resolution by SDS-PAGE. The gels show the peak fractions of GST-A3H haplotype II and GST-A3H haplotype II R44A/Y46A with start and end volumes corresponding to the fractions that were resolved by SDS-PAGE. (G) The Sephadex 200 column (10 mL bed volume) demonstrates that GST is a dimer. The gel shows the peak fractions of GST with start and end volumes corresponding to the fractions that were resolved by SDS-PAGE. However, GST tagged A3H does not dimerize through the GST tag based on data from GST-A3H haplotype II R44A/Y46A (panel D).

in order design mutations that would result in an A3H haplotype II monomer to determine the effect on intersegmental transfer.

Using size exclusion chromatography, we observed that A3H haplotype II and V both primarily form dimers in solution (apparent molecular mass of 44 kDa, Fig. 7.5, A and B). We also observed minor monomer (apparent molecular mass of 24 kDa) and tetramer (apparent molecular mass of 98 kDa) peaks (Fig. 7.5, A and B). A dimer and tetramer of A3H are of equivalent molecular mass to a monomer and dimer of A3G, respectively [140]. The formation of A3H tetramers was not concentration dependent since the chromatograms for 150 μg and 300 μg enzyme showed that the tetramer fraction remained at only 20% (haplotype II) or 25% (haplotype V) of the total population (Fig. 7.5A and inset graph). These data confirmed that D178 was not responsible for mediating dimerization and instead the amino acid at position 178 is a specific determinant that contributes to intersegmental transfer ability. Because helix 6 did not mediate oligomerization for A3H (Fig. 7.5, A and B) as may have occurred based on A3C crystal structure contacts [205], we performed our structure-guided mutagenesis using another single Z-domain oligomerization model, the A2 crystal structure and surrogate mutations made in activation-induced cytidine deaminase [373, 380]. The 40 amino acid N-terminal deleted A2 that was crystallized was found to dimerize through two monomers pairing their β2 strands. Through sequence alignment of A2 and A3H and biochemical data from activation-induced cytidine deaminase [373, 380], we generated an A3H mutant with amino acid substitutions R44A/Y46A located at this predicted oligomerization interface along the β2 strand (Fig. 7.5C) and determined the oligomerization state using size exclusion chromatography.

The A3H haplotype II predicted β2 strand mutant R44A/Y46A was less stable than A3H haplotype II as evidenced by approximately 10-fold less recombinant protein being produced from the *Sf*9 cell system (data not shown). Thus, the GST tag was not cleaved to extend the stability of the mutant for size exclusion chromatography since the R44A/Y46A mutant was prone to precipitation after extended incubation at 4° C in solution without GST. Only 10 μg of GST-A3H wild type and mutant were applied to the column due to the lower yield of GST-A3H haplotype II R44A/Y46A mutant during purification and the chromatograms were generated by manually quantifying integrated gel band intensities in each fraction from an SDS-PAGE gel. Further, the GST-tagged wild type and mutants were analyzed on a 10 mL G200 Superdex

column, in contrast to untagged A3H that was analyzed using a 25 mL G200 Superdex column (see Material and Methods). To ensure that the GST tag did not confound our interpretation of the results, the GST-tagged R44A/Y46A mutant was compared to GST-tagged A3H haplotype II during the size-exclusion chromatography. Due to the 22 kDa GST tag, the predicted GST-A3H monomer molecular weight was 46 kDa. Consistent with β2-β2 strand dimerization, the GST-A3H haplotype II R44A/Y46A mutant eluted predominantly as a monomer (apparent molecular weight 37 kDa, Fig. 7.5, D-F). Although GST alone forms a dimer (apparent molecular weight 51 kDa, Fig. 7.5G), the GST appears to be unable to dimerize when attached to the A3H protein (Fig. 7.5D), otherwise we would be unable to resolve GST-A3H haplotype II R44A/Y46A at a molecular weight of 37 kDa. GST-A3H haplotype II resolved as a dimer (apparent molecular weight 77 kDa, Fig. 7.5, D-F), consistent with the untagged A3H haplotype II data (Fig. 7.5A). There is a clear peak shift of the GST-A3H haplotype II and GST-A3H haplotype II R44A/Y46A in the SDS-PAGE analysis of the fractions suggesting that GST-A3H haplotype II R44A/Y46A fully disrupted dimerization (Fig. 7.5, D and F). Nonetheless, in an attempt to examine the extent of the β2 strand contacts and whether GST-A3H haplotype II R44A/Y46A mutant fully disrupted dimerization we also made a GST-A3H haplotype II P42S mutant, but this mutant was not stably produced from the *Sf*9 cell system. These data support the conclusion that A3H dimerization is mediated by multiple contacts along the predicted β2 strand and involves at least amino acid residues [44]RGY[46]. Further, these data demonstrated that the A3H haplotype II dimer was stable over a 15-fold concentration range (Fig. 7.5, compare A and D).

According to the A2 single Z-domain oligomerization model, a β2-β2 mediated dimer can tetramerize via hydrophobic amino acids on loop 7[373]. A3H haplotype II formed tetramers in solution (Fig. 7.5A). To investigate the amino acids involved in A3H tetramerization we made an A3H haplotype II loop 7 double mutant, Y112A/Y113A (Fig. 7.6A). During size exclusion chromatography of 150 μg of A3H haplotype II Y112A/Y113A predominantly dimers were formed (Fig. 7.6B, apparent molecular weight of 39 kDa), similar to wild type A3H haplotype II (Fig. 7.5A). In contrast to wild type A3H haplotype II, the Y112A/Y113A mutant formed less tetramers (Fig.7.6B, peak denoted by T). At 300 μg of A3H haplotype II Y112A/Y113A, the mutant had no increase in tetramer formation suggesting that Y112 and Y113 primarily

A

Loop 7

```
            112
             |
Hap II  ¹⁰⁹SRLYYHWCKP-¹¹⁸
A2      ¹⁵³GRLF-MWEEPE¹⁶²
```

B

C
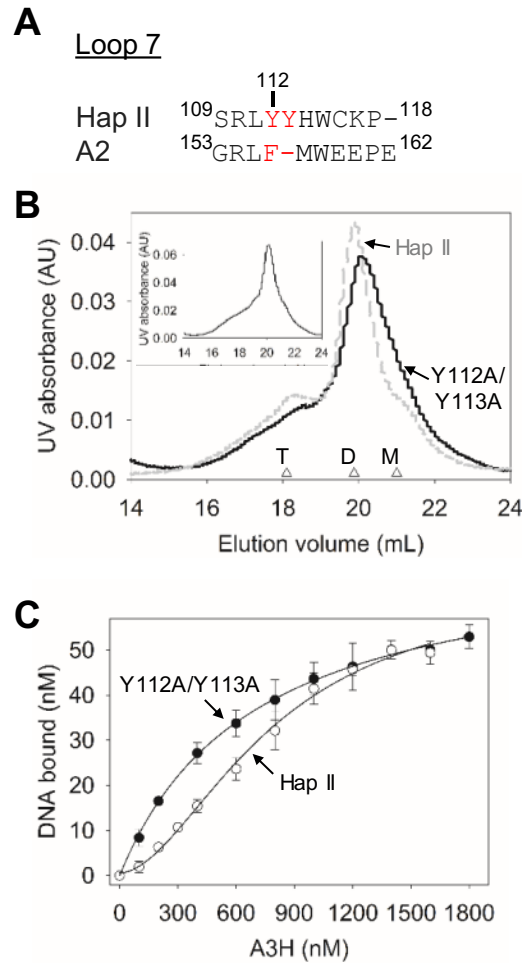
Figure 7.6 Loop 7 amino acids Y112 and Y113 mediate A3H tetramerization. (A) Sequence alignment of A3H haplotype II (Hap II) and APOBEC2 (A2) loop 7 amino acid sequences. Amino acids that were mutated are indicated in red. (B) The size exclusion chromatography profile of A3H Hap II Y112A/Y113A (Y112A/Y113A, 150 μg) was used to calculate the oligomerization state of the enzyme from a standard calibration curve (Fig. 7.5B). Hap II Y112A/Y113A formed dimers (apparent molecular weight of 39 kDa in peak fraction). Overlay: Hap II wild type (dashed line) is shown from Fig. 5A for comparison. A3H Hap II Y112A/Y113A has less of a tetramer peak than haplotype II wild type. Inset graph: A3H Hap II Y112A/Y113A did not resolve with a tetramer peak even when 300 μg enzyme was resolved on the size exclusion column. (C) The apparent K_d of A3H Hap II and A3H Hap II Y112A/Y113A from a 118 nt ssDNA was analyzed by steady-state rotational anisotropy. The A3H Hap II bound ssDNA with a best least squares fit to a sigmoidal binding curve (Hill coefficient of 1.8) and an apparent K_d of 0.72 ± 0.14 μM. The A3H Hap II Y112A/Y113A mutant bound ssDNA with a best least squares fit to a rectangular hyperbola and an apparent K_d of 0.70 ± 0.13 μM.

contribute to tetramerization (Fig. 7.6B, inset graph). However, we could not test this hypothesis further since another loop 7 mutant, A3H haplotype II H114A/W115A, could not be stably produced from *Sf*9 cells, in agreement with the low cellular stability observed by others [213, 371]. All together the data support that A3H haplotype II forms dimers through β2 strand interactions and tetramerizes through loop 7.

In solution the tetramers were a small proportion of the total A3H (Fig. 7.5A). However, A3H haplotype II bound ssDNA cooperatively (Fig. 7.6C, Hill coefficient of 1.8) suggesting that tetramerization was promoted by A3H dimers binding ssDNA. In support of this hypothesis, the A3H haplotype II Y112A/Y113A mutant bound ssDNA with a similar apparent K_d to A3H haplotype II, but the saturation curve fit a noncooperative rectangular hyperbola (Fig. 7.6C). Together these data indicate the formation of tetramers on ssDNA. A3H haplotype II Y112A/Y113A was not enzymatically active on several substrates with different deamination motifs, e.g., 5′ATC, 5′TTC, 5′CTC, 5′GTC, 5′CCC, and 5′AAC, in agreement with a previous study [213, 371] (and data not shown). Loop 7 is known to interact with ssDNA and mediate deamination motif preference in DNA binding models for multiple APOBEC deaminases[184-187]. Since the A3H haplotype II Y112A/Y113A mutant bound ssDNA with an apparent K_d similar to A3H haplotype II the data indicated that the loop 7 mutant was likely inactive due to an altered conformation of ssDNA in the active site (Fig. 7.6C).

7.3.5 A3H haplotype II β2 strand mutant exhibits compromised sliding and intersegmental transfer. We tested A3H haplotype II R44A/Y46A to determine whether this β2 strand mutant was still able to processively scan ssDNA. Interestingly, the A3H haplotype II R44A/Y46A mutant bound ssDNA with a cooperative binding curve (Fig. 7.7A, apparent K_d of 0.21 µM; Hill coefficient of 1.6), suggesting that the mutant can dimerize through loop 7 amino acids on ssDNA and does not remain a monomer as in solution (Fig. 7.5D). Since the Y112A/Y113A mutant was not catalytically active (data not shown), we were unable to test deamination ability of an A3H R44A/Y46A/Y112A/Y113A mutant. Nonetheless, testing the A3H haplotype II R44A/Y46A would reveal the function of the β2-β2 strand dimer on catalytic activity and ssDNA scanning. We found that the A3H haplotype II R44A/Y46A could still processively slide to processively deaminate multiple cytosines, but not as efficiently as A3H haplotype II. On a

143
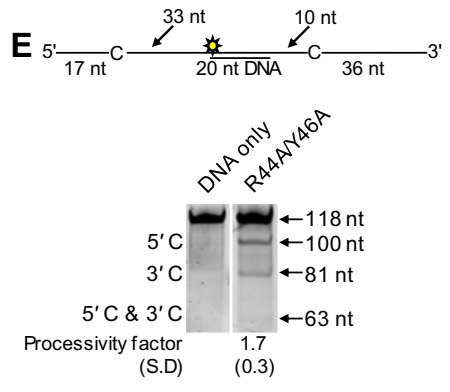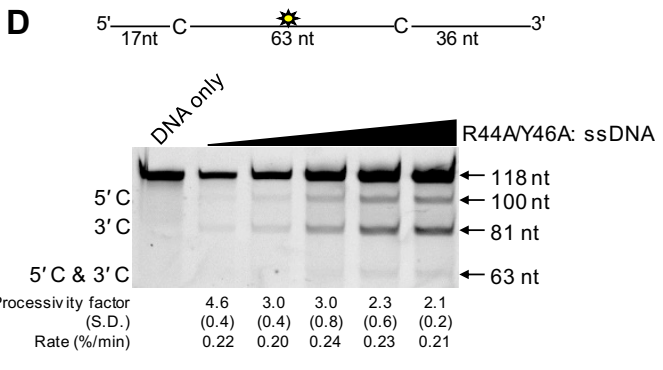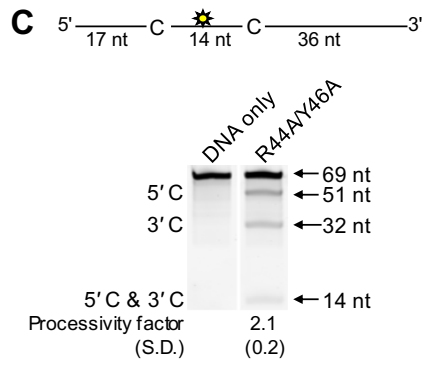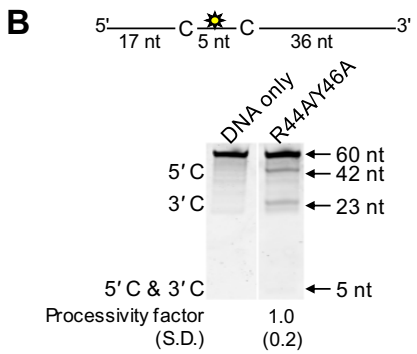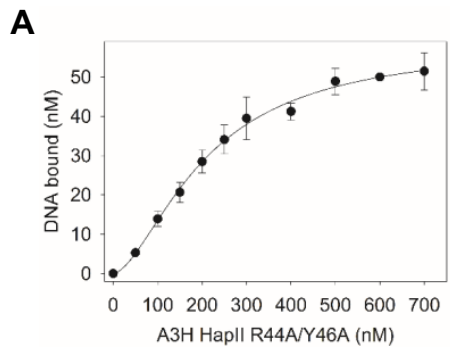
Figure 7.7 Analysis of A3H haplotype II β2 strand mutant R44A/Y46A ssDNA binding and scanning. (A) A3H haplotype II (Hap II) R44A/Y46A ssDNA binding was analyzed by steady-state rotational anisotropy. The apparent dissociation constant (K_d) of A3H Hap II R44A/Y46A from the 118 nt ssDNA was 0.21 ± 0.04 μM. The binding curve had a best least squares fit to a sigmoidal binding curve (Hill coefficient 1.6) demonstrating that the mutant could still oligomerize on ssDNA. (B-E). Processivity of A3H Hap II R44A/Y46A (R44A/Y46A) was tested on ssDNA substrates that contained fluorescein labeled deoxythymidine (yellow star) between two 5'CTC deamination motifs separated by different distances. (B) Deamination of a 60 nt ssDNA substrate with deamination motifs spaced 5 nt apart. Single deaminations of the 5'C and 3'C are detected as the appearance of labeled 42 nt and 23 nt fragments, respectively; double deamination of both C residues on the same molecule results in a 5 nt labeled fragment. (C) Deamination of a 69 nt ssDNA substrate with deamination motifs spaced 14 nt apart. Single deaminations of the 5'C and 3'C are detected as the appearance of labeled 51 nt and 32 nt fragments, respectively; double deaminations of both C residues on the same molecule results in a 14 nt labeled fragment. (D) Deamination of a 118 nt ssDNA substrate with deamination motifs spaced 63 nt apart. Single deaminations of the 5'C and 3'C are detected as the appearance of labeled 100 nt and 81 nt fragments, respectively; double deamination of both C residues on the same molecule results in a 63 nt labeled fragment. For this assay, an A3H: ssDNA ratio of 1:2 was kept constant, but reaction components were increased (50 nM: 100 nM; 100 nM: 200 nM; 150 nM: 300 nM; 200 nM: 400 nM; 250 nM: 500 nM). These reaction conditions enabled the investigation of whether A3H haplotype II R44A/Y46A could translocate on ssDNA by intersegment transfer. Reaction rate of deamination (percent (%)/min) for each individual reaction condition is shown below the gel. (E) Deamination of the same substrate described for (D), but with a 20 nt DNA annealed between the two cytosines. For reactions performed on substrates described for (B-C) and (E), the A3H: ssDNA ratio was 1: 2. The measurements of enzyme processivity (Processivity factor) and the Standard Deviation of the mean (S.D.) are shown below the gels. All values are calculated from at least three independent experiments. Gel images with a white line between lanes indicates that intervening lane(s) on the same gel were cropped for visualization.

ssDNA substrate with cytosine separated by 5 nt, the A3H haplotype II R44A/Y46A had a processivity factor of 1 demonstrating that it was essentially not able to processively deaminate closely spaced cytosines (Fig. 7.7B), in contrast to A3H haplotype II that had a processivity factor 2.1 on this substrate (Fig. 7.1A). When the deamination motifs were spaced further apart, but still within a distance that requires sliding for processive deamination, the A3H haplotype II R44A/Y46A mutant was able to processively deaminate the two motifs (Fig. 7.7C, deamination motifs 14 nt apart, processivity factor of 2.1), but the sliding was not as efficient as A3H haplotype II (Fig. 7.1B, processivity factor of 3.6). These data indicated that although the A3H haplotype II R44A/Y46A mutant had a helix 6 that was able to mediate sliding, the sliding movement was compromised in the absence of the β2-β2 strand dimer. The β2 strand dimerization may form a ssDNA binding groove that promotes processive sliding in combination with helix 6.

To examine if A3H haplotype II R44A/Y46A was still capable of undergoing intersegmental transfer we used the assay that was used for A3H haplotypes II and V (Fig. 7.2). The processivity of A3H haplotype II R44A/Y46A did show a 2-fold decrease in the apparent processivity with an increased concentration of the reaction components (Fig. 7.7D), similar to A3H haplotype II (Fig. 7.2A). However, the rate of the reaction did not increase as the reaction components increased (compare Figs. 7.2A and 7.7D). These data indicate that the A3H haplotype II R44A/Y46A could transfer to other ssDNA substrates, likely due to the dimerization mediated by loop 7 (Fig. 7.7A, cooperative binding curve), but did not readily cycle through the ssDNA substrates (Fig. 7.7D, no increase in reaction rate). This may be due to the loop 7 dimer being less stable than the β2 strand dimer, which is supported by the lack of a concentration dependence of tetramer formation in solution (Fig. 7.5A). However, the data suggested that A3H haplotype II R44A/Y46A mutant was able to undergo 3-D translocations since the processivity on the ssDNA substrate with the deamination motifs separated by 63 nt was similar to that of A3H haplotype II (compare processivity factors of 5.2 (Fig. 7.1C) and 4.6 (Fig. 7.7D, first reaction lane)). The 3-D translocation ability of A3H haplotype II R44A/Y46A was confirmed by demonstrating that the mutant retained processivity in the presence of a dsDNA region between the deamination motifs (Fig. 7.7E).

146

7.4 Discussion

Human A3H is the most polymorphic human A3 enzyme with seven different haplotypes having different cellular stabilities [45, 195, 198]. In comparison to other A3s, stable A3H haplotypes are most able to have a restrictive effect on the progression of an HIV infection in the early stages [100, 199, 232]. This is due to A3H not being sensitive to all of the HIV Vif variants, unlike A3G and A3F [45, 211, 232, 240, 276]. Because the majority of humans of non-African descent carry A3H alleles that code for an unstable form of A3H, not all HIV Vif variants are adapted to induce degradation of A3H if they are circulating in a geographical location where the majority of the population carries unstable A3H alleles [198, 276]. Thus, if an HIV strain from a human with inactive A3H alleles infects a human with a stable A3H allele, the Vif will have to adapt to degrade the stable A3H [100, 232]. It has been found that stable A3H can suppress HIV replication and delay time to treatment early in infection [100]. Thus, stable A3H alleles could potentially act as an HIV infection barrier [232].

Despite the realized importance of A3H in the A3 repertoire that can suppress HIV replication, there has been no biochemical characterization of the ssDNA scanning mechanisms of A3H, a biochemical comparison of stable haplotypes, or identification of the residues mediating oligomerization. Because oligomerization is required for virion encapsidation and enables A3H to be the only single Z-domain A3 naturally capable of restricting HIV replication [78, 191], understanding the interfaces where oligomerization occurs is important for realizing the determinants for HIV restriction. Thus, we focused our study on the two stable haplotypes, II and V, which occur with the highest frequency in humans [45, 198]. Despite only differing by a single amino acid, we found biochemical differences in haplotypes II and V. The more rigid side-chain and lower isoelectric point of the A3H haplotype II D178 in comparison to the A3H haplotype V E178 appeared to affect the ssDNA scanning mechanisms and the binding affinity for dsDNA and RNA. Our study of A3H has established the basis of processivity for a single Z-domain A3 enzyme, that redundant processive mechanisms are used by A3H, a function of the A3H amino acid 178 polymorphism, and that helix 6 influences dsDNA and RNA binding affinity.

In general, an enzyme that moves processively on DNA using intersegmental transfer must have at least two DNA binding domains to mediate the doubly-bound intermediate state where the enzyme simultaneously binds two segments of DNA before transferring to one of the

bound segments [146, 148]. However, our data show that intersegmental transfer ability can be mediated by A3H at different levels. Both A3H haplotype V and A3H haplotype II R44A/Y46A have altered intersegmental transfer ability, but for different reasons (Figs. 7.2B and 7.7D). The A3H haplotype II and V data (Fig. 7.2) demonstrates that D178 on predicted helix 6 promotes enzyme cycling during intersegmental transfer. The absence of D178, either by mutating A3H to a K178 or studying A3H haplotype V that has an E178 inhibited the increased reaction rate associated with intersegmental transfer despite the presence of two DNA binding domains (compare Figs. 7.2 (A and B), 7.4B, and 7.5A). This data is consistent with Mitra *et al.* [213], who found helix 6 was important for A3H activity and ssDNA binding, although they did not test processivity. The polymorphism at amino acid position 178 has no other known function [197]. The A3H haplotype II R44A/Y46A mutant demonstrated that the dimer interface is also important for promoting intersegmental transfer. Despite cooperative binding of A3H haplotype II R44A/Y46A to ssDNA, signifying the formation of oligomers, likely through loop 7, the mutant was unable to cycle efficiently after transfers to another DNA molecule (Fig. 7.7, A and D). This may either be due to the loop 7-mediated dimer being less stable than a β2 strand-mediated dimer or that the β2 strand dimer interface creates a ssDNA binding groove along the enzyme that promotes interaction of the ssDNA with helix 6. Intersegmental transfer ability is unique to A3H among the A3 enzymes characterized thus far [133]. Although A3G has been characterized to undergo intersegmental transfer in one report [93], this has not been confirmed by others [133, 149, 381]. Further, testing A3G for intersegmental transfer movement under the same reaction conditions as A3H showed that the processivity of A3G did not decrease with increasing concentration of reaction components (data not shown), demonstrating that A3G uses primarily jumping as a 3-D translocation mechanism [132]. It is not known why A3 enzymes such as A3G or A3F do not readily undergo intersegmental transfer since there are multiple DNA binding domains within a monomeric unit [37, 133, 139]. Based on our results with A3H, this may be due to differences in the ssDNA binding properties of the helix 6. Based on results with A3C, another factor may be the flexibility between the two ssDNA binding domains [216]. Similar to A3G and A3F, A3H also uses helix 6 to mediate sliding movements [133, 139]. Jumping movements of A3G are mediated by loop 7 [139], however, we could not test this with A3H since the loop 7 mutant was not catalytically active (data not shown), in agreement with previous studies [198, 213].

A3H haplotype II was found to use several processive mechanisms to scan ssDNA. A3H haplotype II could slide on ssDNA, slide over RNA/DNA hybrids, and translocate by intersegmental transfer and jumping (Fig. 7.1 and 7.2A). Although A3H haplotype V exhibited diminished sliding and intersegmental transfer ability (Fig. 7.1 A-B and 7.2B), the enzyme was able to jump (Fig. 7.1, C- D) and restricted HIV replication similarly to A3H haplotype II (Fig. 7.3). Thus, the data indicate that for A3H the 3-D translocation mechanisms of intersegmental transfer and jumping are redundant. A3H haplotype V could processively slide on ssDNA to deaminate cytosines that were 14 nt apart but not 5 nt apart (Fig. 7.1, A-B). However, A3H haplotype V-mediated mutagenesis of HIV was not adversely affected, despite a lack of a redundant mechanism for searching for closely spaced cytosines (Fig. 7.3). These data suggest that deficiencies in sliding are not as detrimental to the efficiency of the search for deamination motifs as deficiencies in jumping. This is supported by A3G reaction conditions and A3G mutants that removed either sliding or jumping ability [139, 183]. For A3G, larger decreases in mutagenic efficiency result from losses in jumping rather than sliding [139, 183]. Altogether, the A3H data are consistent with our previous work which found that a high A3 mutagenic potential during HIV reverse transcription requires the enzyme to use both a 1-D and 3-D translocation mechanism [133]. Our data are also in agreement with others that have shown A3H haplotype V to be as restrictive as A3H haplotype II in primary cells during a spreading infection and in single-cycle infectivity assays [198, 211, 212, 232].

Our results indicate that A3H oligomerizes similar to the model proposed for the N-terminal 40 amino acid deleted A2 crystal structure [373]. The A3H dimer is highly stable and even at very low concentrations of A3H, the dimer interface was maintained (Fig. 7.5D, 10 μg). The A2 model of oligomerization where a single Z-domain enzyme dimerizes through the β2 strands and tetramerizes through loop 7 has been questioned since this structure only forms in the N-terminal deleted version of A2 and human A2 *in vitro* activity has yet to be demonstrated [191, 373, 382]. The A3H data demonstrates that the A2 model of oligomerization is a valid model for single Z-domain enzymes. A caveat of our data is that we could not produce a β2 strand mutant that was as stable as wild type A3H haplotype II. Thus, we required a GST tag to stabilize the β2 strand mutant. It is known that GST alone forms dimers (Fig. 7.5G), but our data demonstrate that GST-A3H does not form dimers through the GST tag or we would be unable to resolve A3H haplotype II R44A/Y46A as a monomer (Fig. 7.5D). These data suggest that the A3H can block

GST dimerization in our GST-A3H fusion construct. Nonetheless, this model of oligomerization does not appear to be conserved across all Z-type domains since multiple crystal structures of single Z-domain enzymes or C-terminal domains of double Z-domain enzymes have not oligomerized according to the A2 model [205, 365, 383-385]. Because A3H is the only Z3-domain in humans it may be the only A3 to oligomerize in this manner [37]. A3H tetramers form primarily during DNA binding and not in solution. This is supported by the data for the A3H haplotype II β2 strand mutant R44A/Y46A that could still bind ssDNA cooperatively (Fig. 7.7A) and the A3H haplotype II loop 7 mutant Y112A/Y113A that could not (Fig. 7.6C). These data are also consistent with size exclusion chromatography profiles of A3H haplotypes II and V where tetramers were a minority of the population in comparison to dimers (Fig. 7.5A). Mitra *et al*. [213] have investigated the role of loop 7 in the A3H-ssDNA interactions required for catalytic activity. Deamination reactions conducted with A3H haplotype II and mutants expressed from 293T cells demonstrated that Y112A and W115A mutants lose deamination activity [213]. Interestingly, W115F rescued the deamination activity by 30%, suggesting that aromatic residues on loop 7 are involved in nucleic acid interactions that are required for catalysis through base stacking [213], in agreement with data from activation-induced cytidine deaminase [386] . However, this does not affect the apparent K_d from ssDNA (Fig. 7.6C).

The A3H dimer organization may provide an advantage by decreasing the Vif-mediated degradation efficiency and thus promoting the HIV restriction ability of A3H. We previously showed that A3G dimerization is disrupted by Vif [312]. This may contribute to efficient Vif-induced polyubiquitination. Vif can disrupt A3G dimerization since the primary residue Vif interacts with on A3G is Asp-128, which is adjacent to the primary dimerization amino acids, Phe-126 and Trp-127 [64, 138, 278, 279, 312]. In contrast, when A3H haplotype II interacted with Vif, there was a structural change induced by Vif, but this did not appear to be a complete disruption of dimerization [312]. Our data on the A3H dimer interface suggest that this is because A3H dimerizes through the β2 strands (Fig. 7.5D) and Vif interacts with A3H on a separate structure, predicted helix 4 [265, 387]. Using quantitative immunoblotting it was shown that Vif_{HXB2} is 4-fold less efficient at degrading A3H than A3G. Further, both Vif_{NL4-3} and Vif_{HXB2} could induce polyubiquitination of A3H with K63 and K48 linkages, in contrast to A3G that was polyubiquitinated with only K48 linkages [312]. Thus, the A3H β2 strand dimer may decrease the effectiveness of Vif and promote alternate ubiquitination modifications. Although during

viral infections, HIV Vif can effectively suppress both A3G and A3H to enable viral replication [276], the higher resistance of A3H to Vif may provide an advantage during the early stages of infection and during Vif adaptation [100, 232].

7.5 Material and Methods

7.5.1 Cloning and site-directed mutagenesis. An A3H haplotype I clone (NCBI Accession BC069023) was obtained from Open Biosystems and site directed mutagenesis was used make A3H haplotype II (G105R, K121D, E178D) and haplotype V (G105R, K121D). A GST-A3H construct was subcloned from a pAcG2T vector (BD Biosciences) into a pFAST-bac1 vector (Life Technologies) using EcoRI and NotI restriction sites. Mutants of A3H haplotype II (R44A/Y46A, Y111A/Y112A, D178K, and D178A) were made using the A3H haplotype II construct for site-directed mutagenesis (QuickChange site-directed mutagenesis protocol, Stratagene). The cDNA for A3H haplotypes II and V was also cloned into pcDNA3.1 with an HA tag. For generation of RNA *in vitro*, the HIV 5'UTR (nucleotides 1-497) and human Alu sequence were cloned into pSP72 vector (Promega) using XhoI and HindIII sites under the control of T7 promoter (5'UTR) or SP6 promoter (Alu). All constructed plasmids were verified by DNA sequencing. Primers were obtained from Integrated DNA Technologies and are listed in Table 7.5.

7.5.2 Protein expression and purification. The pFASTbac1-GST-A3H vectors were used to produce recombinant baculovirus according to the protocol for the Bac-to-Bac system (Life Technologies). Recombinant GST-A3H baculovirus were then used to infect *Sf*9 cells at a Multiplicity of Infection (M.O.I) of 20 and cells were harvested after 72 h. Cells were lysed and clarified lysates were incubated with Glutathione Sepharose 4B resin (GE Healthcare) at 4°C and subjected to a series of salt washes as previously described. For all A3H haplotypes and mutants, except the A3H haplotype II R44A/Y46A, D178K, and D178A mutants, on-column cleavage from the GST tag with thrombin (GE Healthcare) was performed at 21°C for 18 h in thrombin digestion buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, and 1mM DTT). Enzymes were assessed to be 95% pure by SDS-PAGE. Due to the lower expression of the A3H haplotype II R44A/Y46A, D178K, and D178A mutants, the enzyme was eluted with the GST tag in elution buffer (100 mM Tris pH 8.8, 150 mM NaCl, 10% (v/v) glycerol and 50 mM reduced glutathione). A wild type GST-haplotype II was purified in parallel with the mutant for comparative analyses. Eluted GST-A3H was dialyzed against 100 mM Tris pH 7.5, 250 mM

Table 7.5. Primers and DNA substrates

Name	Sequence
pFASTbac1-A3H Forward	GAA TTC ATG TCC CCT ATA CTA GGT TA
pFASTbac1-A3H Reverse	GCG GCC GCT CAG GAC TTT ATC CTC TC
G105R SDM Forward	GCT CAC GAC CAT CTG AAC CTG CGC ATC TTC GCC TCC CGC CTG
G105R SDM Reverse	CAG GCG GGA GGC GAA GAT GCG CAG GTT CAG ATG GTC GTG AGC
K121D SDM Forward	TGC AAG CCC CAG CAG GAC GGG CTG CGG CTT CTG
K121D SDM Reverse	CAG AAG CCG CAG CCC GTC CTG CTG GGG CTT GCA
E178D SDM Forward	ATA AAG CGA CGG CTT GAC AGG ATA AAG TCC TGA
E178D SDM Reverse	TCA GGA CTT TAT CCT GTC AAG CCG TCG CTT TAT
Y112A/Y113A SDM Forward	ATC TTC GCC TCC CGC CTG GCC GCT CAC TGG TGC AAG CCC CAG
Y112A/Y113A SDM Reverse	CTG GGG CTT GCA CCA GTG AGC GGC CAG GCG GGA GGC GAA GAT
H114A/Y115A SDM Forward	GCC TCC CGC CTG TAC TAC GCC GCT TGC AAG CCC CAG CAG GAC
H114A/Y115A SDM Reverse	GTC CTG CTG GGG CTT GCA AGC GGC GTA GTA CAG GCG GGA GGC
R44A/Y46A SDM Forward	AAT GGC TCC ACG CCC ACG GCC GGC GCT TTT GAA AAC AAG AAA AAG
R44A/Y46A SDM Reverse	TTT TTC TTG TTT TCA AAA GCG CCG GCC GTG GGC GTG GAG CCA TT
D178K SDM Forward	ATA AAG CGA CGG CTT AAG AGG ATA AAG TCC TGA
D178K SDM Reverse	TCA GGA CTT TAT CCT CTT AAG CCG TCG CTT TAT
D178A SDM Forward	ATA AAG CGA CGG CTT GCT AGG ATA AAG TCC TGA
D178A SDM Reverse	TCA GGA CTT TAT CCT AGC AAG CCG TCG CTT TAT
118 nt (CTC) ssDNA (deamination	GAA TAT AGT TTT TAG CTC AAA GTA AGT GAA GAT AAT [Fam-dT] TAG AGA GTT GTA ATG TGA TAT ATG TGT ATG AAA GAT ATA AGA CTC

assay and anisotropy)	AAA GTG AAA AGT TGT TAA TGT GTG TAG ATA TGT TAA
85 nt (CTC) ssDNA	AAA GTG AAA GTG ATA CTC AAA TTT AAA AGT [Fam-dT] AGA TAG AAG GTG ATA CTC AAA TAT GAA AGT TAG TAA GAT GTG TAA GTA TGT TAA
20 nt DNA	CTT TCA TAC ACA TAT ATC AC
20 nt DNA complement	Fam-GTG ATA TAT GTG TAT GAA AG
20 nt RNA	CUU UCA UAC ACA UAU AUC AC
20 nt RNA complement	Fam-GUG AUA UAU GUG UAU GAA AG
69 nt (CTC) ssDNA	AAG AAG TTA GTA AGA CTC AAA ATT [Fam-dT] AAA AAC TCA AAT GTT AGA TAT GTT AAT GTG TGT GAT GAT GTT GA
60 nt (CTC) ssDNA	AAA GTA AAA GTG ATA CTC A [Fam-dT] GCT CAT AGA TTA AAG TTA GTA AGA TGT GTA AGT ATG TTA A
prot Forward	GAC AAG GAA CTG TAT CCT TTA GCT T
prot Reverse	CTG GTA CAG TCT CAA TAG GAC TAA T
HIV 5'UTR Forward	GAA TTC GCC AGA GAG CTC CCG GGC
HIV 5'UTR Reverse	AGA TCT TGG AAG GGC TAG TTT ACT C

Fam denotes Fluorescein amidite

NaCl, 10% glycerol, and 1mM DTT overnight at 4°C. For size exclusion chromatography with the A3H haplotype II R44A/Y46A mutant, the GST tag was not cleaved to increase protein stability. For deamination reactions with the A3H haplotype II R44A/Y46A, D178K, and D178A mutants the GST tag was cleaved in solution with thrombin (GE Healthcare).

7.5.3 Single-cycle infectivity assay. VSV-G pseudotyped HIV pNL4-3 *ΔVif* viruses were produced by transfecting 3×10^5 293T cells (ATCC CRL-3216) per well in a 6-well plate with GeneJuice (Novagen). Specifically, transfections used 500 ng of pHIV*ΔVif*, which expresses an eGFP reporter gene and 200 ng of pMDG, which expresses the VSV-G protein, in the presence or absence of A3H haplotype II or A3H haplotype V in pcDNA3.1. Empty pcDNA3.1 was used to achieve equivalent amounts of transfected DNA. A titration of A3H expression vector was used that ranged from 25-200 ng. The amino acid sequences of the A3s were identical to those used in biochemical assays. Twenty-four hours after the transfection the medium was replaced. Virus-containing supernatants were collected 48 hours after the media change and filtered through 0.45 μm syringe filters. Virus was quantified by a p24 enzyme-linked immunosorbent assay (QuickTiter Lentivirus Titer Kit, Cell Biolabs Inc. or HIV-1 p24 ELISA Kit, XpressBio). Target 293T cells were infected at an MOI of 0.5 by spinoculation at $800 \times g$ for 1 h in the presence of 8 μg/ml of polybrene. Infection levels in 293T cells was determined by flow cytometry by detecting eGFP fluorescence at 40 hours post infection and data were normalized to HIV*Δvif* infections in the absence of A3 enzymes.

7.5.4 Immunoblotting. To detect A3H protein levels, HA-tagged versions of the A3H enzymes used for the single-cycle infectivity assays were constructed in pcDNA3.1. For detection of A3H in cell lysates (50 μg total protein) and virions (50 ng of p24) anti-rabbit HA (Sigma) was used. Loading controls for cell lysates (α-tubulin, Sigma) and virions (p24, Cat #3537, NIH AIDS Reagent Program) were detected using mouse monoclonal antibodies. Tagged proteins and loading controls were detected simultaneously on the same immunoblot by using the Licor/Odyssey system (IRDye 680-labeled goat anti-rabbit and IRDye 800-labeled goat anti-mouse secondary antibodies). Antibodies were used at a dilution of 1/1000.

7.5.5 Sequencing of integrated proviral DNA. Infected 293T cells were harvested after 40 h and the DNA was extracted using DNAzol Reagent (Life Technologies). DNA was treated with DpnI (New England Biolabs) for 1 hour at 37 °C to remove possible contaminating plasmid DNA and

the *prot* (nt 2280-2631) sequences were amplified by PCR using Q5 Polymerase (New England Biolabs). Primers were obtained from Integrated DNA Technologies and are listed in Table 7.5. PCR products were purified and cloned with the Zero Blunt TOPO PCR cloning kit (Life Technologies) or NEB PCR Cloning Kit (New England Biolabs). DNA was sequenced with kit specific primers and carried out at the National Research Council of Canada (Saskatoon, Saskatchewan) or Eurofins Genomics (Huntsville, Alabama).

7.5.6 In vitro deamination assay. All ssDNA substrates were obtained from Tri-Link Biotechnologies and are listed in Table 7.5. All reactions were carried out under single-hit conditions, i.e., <15% substrate usage, to ensure that deaminations on each ssDNA were catalyzed by a single enzyme. Under these conditions, a processivity factor can be determined by comparing the total amount of deaminations occurring at two sites on same ssDNA to a calculated theoretical value of deaminations at these two sites if the deamination events were uncorrelated (not processive). In order to achieve less than 15% substrate usage, the ssDNA substrate containing two 5'CTC motifs (100 nM) was incubated with 50 nM of A3H for 1.5 to 20 min at 37 °C in deamination buffer containing 50 mM Tris pH 7.5, 40 mM KCl, 10 mM $MgCl_2$, and 1 mM DTT. The reaction time was varied according to the specific activity of the enzyme on each ssDNA to achieve ~10% substrate usage. DNA substrates with a dsDNA or RNA/DNA hybrid were formed by heat annealing in 50 mM Tris pH 7.5 and 50 mM NaCl. Reactions were started by the addition of the ssDNA substrate. To test if intersegmental transfer was occurring, the A3H: ssDNA ratio (1:2) was kept constant, but increased amounts of enzyme and substrate were titrated into the reaction (ssDNA, 100 nM-700 nM; A3H, 50 nM-350 nM). A3H catalyzed deaminations were detected by treating the ssDNA with Uracil DNA Glycosylase (New England Biolabs) and heating under alkaline conditions before resolving the Fluorescein labeled ssDNA on 10%, 16% or 20% (v/v) denaturing polyacrylamide gels, depending on the sizes of the ssDNA fragments. Gel photos were obtained using a Typhoon Trio multipurpose scanner (GE Healthcare) and integrated gel band intensities were analyzed using ImageQuant (GE Healthcare) as previously described [132].

7.5.7 Size exclusion chromatography. The oligomerization states of A3H haplotype II, haplotype V, and haplotype II Y112A/Y113A were determined by loading 150 or 300 μg of the purified enzymes onto a Superdex 200 10/300 Increase (GE Healthcare). The running buffer contained 20 mM Tris pH 8.0, 300 mM NaCl, 10% (v/v) glycerol and 1 mM DTT. The oligomerization states

of GST-A3H haplotype II and GST-A3H haplotype II R44A/Y46A were examined by loading 10 μg of the purified enzymes onto a size exclusion chromatography column. The column was prepared by pouring a 10 mL Superdex 200 (GE Healthcare) resin bed contained in a column with a 0.5 cm diameter and 16 cm height. The running buffer contained 20 mM Tris pH 8.0, 300 mM NaCl, 10% (v/v) glycerol and 5 mM DTT. The Bio-Rad gel filtration standard set was used to generate a calibration curve from which the apparent molecular masses and oligomerization states of the enzymes were determined.

7.5.8 Steady-state rotational anisotropy. The apparent K_d of A3G, A3H haplotype II, and A3H haplotype V for ssDNA, dsDNA, RNA/DNA hybrid, or RNA were determined using steady state fluorescence depolarization (rotational anisotropy). The ssDNA substrate was the 118 nt ssDNA used for deamination assays and contained an internal Fluorescein-dT (Table 7.5). The dsDNA and RNA/DNA hybrid were composed of 20 nt that matched the hybrid regions used to block A3H sliding in deamination assays (Table 7.5).

Fluorescently labeled RNA was produced by transcribing linearized pSP72 DNA (cut with EcoRI (5'UTR) or HindIII (Alu)) *in vitro* using either T7 RNA polymerase (5'UTR, Roche Life Sciences) or SP6 polymerase (Alu, Roche Life Sciences) with a nucleotide mixture containing fluorescein-12-UTP (Roche).

Reactions (50 μl) were conducted in deamination buffer and contained 50 nM fluorescein labeled DNA or RNA and increasing amounts of A3 (A3G, 0-1000 nM; A3H haplotypes II and V, 0-1500 nM). A QuantaMaster QM-4 spectrofluorometer (Photon Technology International) with a dual emission channel was used to collect data and calculate anisotropy. Measurements were performed at 21°C. Samples were excited with vertically polarized light at 495 nm (6 nm band pass) and vertical and horizontal emissions were measured at 520 nm (6 nm band pass). The K_d was obtained by fitting to a rectangular hyperbola or sigmoidal curve equation using SigmaPlot 11.2 software.

7.6 Acknowledgements

8.0 Exploration of the biological activities of A3 enzymes apart from HIV restriction

Aside from the antiviral activities against exogenous viruses such as HIV, several studies suggest the primitive function of A3 proteins is to prevent the propagation of endogenous retroelements [388, 40]. A3s likely represent an ancient form of antiviral defense in mammals because *A3* genes have been undergoing positive selection throughout the past 30 million years of primate evolution and their existence predates HIV [317]. Endogenous retroelements, including the Long Terminal Repeat (LTR) and non-LTR retrotransposons, represent a class of mobile genetic elements that propagate through an RNA intermediate and they share with retroviruses a reverse transcription step (reviewed in [315, 321]). Uncontrolled retrotransposon mobilization poses a deleterious effect on the host genome stability and A3 proteins have been reported to act differentially to suppress the activities of various retrotransposons [35, 165, 325, 332, 334-336, 389-391]. In Chapter 7, having comparatively characterized the two most frequently occurring A3H haplotypes (Hap II and Hap V) and having elucidated the biochemical determinants in A3H that are required for HIV restriction, we next set out to examine if the same determinants are needed for restriction of retroelements. Specifically we determined how two A3H haplotypes restrict the replication of the non-LTR retrotransposon L1, the only active autonomous retrotransposable element present in humans (reviewed in [315, 321]). Although A3H's effects on L1 retrotranspositions have been reported in the literature [45, 289], A3H polymorphism confounded the analysis and no studies have clarified the influence of different stable haplotypes on L1 replication. Furthermore, despite A3H Hap II and Hap V restricting HIV with similar efficiencies, we found the dsDNA and RNA binding abilities differ greatly between the two haplotypes [210]. This finding prompted us to test whether the difference in nucleic acid binding between A3H Hap II and Hap V leads to a difference in their abilities to affect L1 replication. In Chapter 9, the L1 restriction capacities of A3H Hap II and Hap V have been characterized and compared to that of A3A and A3G, two A3s that have been shown to be either a robust L1 restrictor or unable to restrict L1 mobilizations, respectively [41, 333, 334, 390, 391].

The experiments performed in Chapter 9 were conceived by Y.F and L.C and were carried out by Y.F with the assistance from M.G. Y.F, M.G and L.C analyzed the data, and Y.F and L.C wrote the paper.

9.0 The inhibitory mechanisms employed by A3A, A3G and A3H against human retrotransposon LINE1

Yuqing Feng[1], Mariam Goubran[1] and Linda Chelico[1]
[1]Department of Microbiology & Immunology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada, S7N 5E5

9.1 Abstract

L1 is the only autonomous non-LTR retroelement currently present in humans and the uncontrolled mobilizations of L1 have been associated with various hereditary diseases and represent a threat to genome stability. A3 proteins have been reported to act differentially against L1 retrotranspositions, although the mechanisms of A3-mediated L1 inhibition have not been well understood. The prevailing view for many years was that A3-mediated L1 inhibition was deamination independent and relied on A3s blocking the L1 reverse transcriptase DNA polymerization. However, during L1 replication, A3s may act on transiently exposed single-stranded cDNA to catalyze the formation of uracils which are excised by the uracil DNA N-glycosylase (UNG), which generates abasic sites and may lead to cDNA degradation. To test the hypothesis that A3-deamination mediated L1 restriction was occurring we included a uracil DNA glycosylase inhibitor (UGI) in experiments that would inhibit UNG and prevent the degradation of L1 cDNA. We predicted that this would allow the recovery of the A3s-catalyzed deamination events. By utilizing a neo[R]-based L1 retrotransposition assay and inspecting A3-mediated G-to-A deamination events on transposed *neo* sequence, we confirmed A3A is a potent inhibitor of L1 retrotransposition, whereas A3G has very little inhibitory effect against L1. We also reported, for the first time, that two most frequently occurring stable A3H haplotypes (Hap II and Hap V) restrict L1 activity using a deamination-independent mechanism(s), in contrast to A3A. Our study demonstrates different A3 proteins have been evolved to inhibit L1 retrotransposition activities through distinct mechanisms.

9.2 Introduction

Transposable elements are DNA sequences that can move from one location of the genome to another. Approximately 45% of the human genome is recognized as being derived from transposable elements. Transposable elements can be grouped into two major classes based on whether they mobilize via a DNA intermediate (i.e., DNA transposons) or an RNA intermediate (i.e., retrotransposons) [320, 321]. DNA transposons, although not being able to

mobilize in the current human genome, have played a major role during the evolution of the eukaryotic genome [37, 392, 393]. Retrotransposons transpose through RNA intermediates followed by reverse transcription and integration at new genomic locations. Retrotransposons can be divided into two classes: Long Terminal Repeat (LTR) retrotransposons (also known as endogenous retroviruses) and the non-LTR retrotransposons. The genetic construct and replication strategy of endogenous retroviruses resembles those of the retroviruses, except the endogenous retrovirus lacks a functional *envelope (env)* gene and their life cycles are strictly intracellular [320, 321].

Non-LTR retrotransposons are found throughout the eukaryotes, and examples of non-LTR retrotransposons include LINE1 (Long Interspersed Nuclear Element) and SINE (Short Interspersed Nuclear Element). Human L1s account for ~17% of the genomic DNA with ~500,000 copies identified, and they represent the only autonomous transposable elements that are currently active in humans [324, 394]. Although the vast majority of L1s are inactive due to accumulated mutations or 5' truncations, human genome still contains 80~100 copies of intact L1s that are retrotransposition competent [34, 315, 320, 321]. Full-length L1 contains a 5' UTR, two Open Reading Frames (ORFs) that are separated by a 63 nt inter-ORF region, and a 3' UTR that ends with a polyadenosine-rich sequence (poly(A) tail) [395]. Proteins encoded by ORF1 and ORF2 are both required for efficient L1 retrotransposition. ORF1 encodes a 40kDa nucleic binding protein with demonstrated nucleic acid chaperone activity; ORF2 encodes proteins with endonuclease (EN) and reverse transcriptase (RT) activities [224, 396, 397]. L1-encoded EN preferentially cleaves ssDNA at an AT-rich consensus sequence (5' TTTT/A, where "/" dictates the cleavage sites). L1 RT is a DNA/RNA dependent DNA polymerase and it lacks a detectable RNaseH activity [398], a trait that distinguishes itself from the retroviral RTs. As a result, L1 RNA could be either degraded by a yet unidentified cellular protein or displaced during (+) strand cDNA synthesis [399], although experiments are required to test these hypotheses. Cellular RNaseH2 cooperates with flap endonuclease 1 (FEN-1) to remove RNA primers during lagging strand DNA synthesis and serves as a good candidate [400]. Unlike retroviruses and LTR retrotransposons which undergo reverse transcription in cytoplasm, L1 RNA is reverse-transcribed in the nucleus [401, 402] (Fig. 1.8B). L1 relies on a mechanism termed the Target Primed Reverse Transcription (TPRT) for retrotransposition [403] (Fig. 1.9).

159

Despite the positive effects of retroelement mobilizations, such as contributing to genome evolution and gene diversity, excessive transposition events need to be suppressed, as several genetic diseases have been associated with L1-mediated insertional mutagenesis in germline and somatic cells [34, 315]. The retroelement suppression mechanisms include DNA methylation [404, 405], RNA interference [406, 407], cellular DNA repair factors [408, 409]. Also, recent studies suggest that several cellular proteins are involved in L1 restriction. Examples of these cellular proteins include Three primer repair Exonuclease 1 (TREX1) [410], Zinc finger Antiviral Protein (ZAP) [411], RNA helicase Mov10 [412], RNaseL [413], SAM domain and HD domain-containing protein 1(SAMHD1) [414, 415], and APOBEC family of cytosine deaminases [35, 165, 325, 332, 334, 335, 389-391].

APOBECs emerged at the origin of vertebrates and have diversified throughout the vertebrate lineage [37, 317]. The most ancient APOBEC family members include Activation Induced cytidine Deaminase (AID) and APOBEC2 (A2) proteins that are identified in jawless and cartilaginous fish [37]. AID is involved in class-switch recombination and somatic hypermutations in adaptive immune response and A2 proteins are involved in cardiac and skeletal muscle development [416, 417]. APOBEC4 (A4) and APOBEC5 (A5) then emerged in amphibians, although their functions are yet to be identified [418, 419]. Throughout tetrapod evolution, duplication of AID leads to the emergence of APOBEC1 (A1) [420]. In placental mammals the A3 locus evolved and greatly expanded [37]. While only one *A3* gene is present in mice, up to seven *A3* (*A3A-A3H*) genes are present in primates. *A3* genes have likely acquired the anti-L1 activities from the *AID/A2*-like gene from which they have evolved. Several studies have clearly demonstrated that in cell culture system, A3 family members could possess inhibitory activities against a number of mammalian L1s [35, 165, 325, 332, 334, 335, 389-391]. However, which cellular factors control the mobility of L1 in species that lack *A3* genes became an intriguing question and it seemed reasonable to hypothesize that L1 restriction by APOBEC proteins predates mammals, as *A3* genes are only limited to mammalian lineages, whereas the L1s are far more ancient, dating back to the emergence of vertebrates. It was not surprising to find that APOBEC proteins, such as pre-mammalian AID and reptilian A1 from green anole lizard, are indeed capable of inhibiting non-mammalian L1 genomes [35, 388].

A3s have been reported to restrict L1 by both deamination-dependent or -independent mechanisms [41, 325, 332-335, 391, 421]. For a deamination-dependent mechanism, during the

reverse transcription of the L1 RNA, (-)cDNA gets transiently exposed after the RNA template degradation, which renders it susceptible to A3-mediated C→U deaminations. The uracils present on L1 cDNA could trigger cDNA degradation through the actions of host base excision repair pathway. Specifically, cellular uracil DNA N-glycosylase (UNG) proteins excise the uracils from ssDNA, generating abasic sites and apurinic/apyrimidinic endonuclease (APE) produces nicks on DNA which would induce L1 cDNA degradation. Alternatively, an uracil-containing cDNA may escape the degradation, becomes replicated to form a double-stranded (ds) DNA and integrates into the host chromosome and is functionally inactivated through A3 induced C/G→T/A transition mutations. For a deamination-independent mechanism of L1 restriction, the A3s can sequester L1 ribonucleoprotein (RNP) complex to the cytoplasmic compartment such as stress granule or physically interact with L1 RT and inhibit DNA polymerization during TPRT [42, 335, 336]. There are controversies in the literature as to whether A3s utilize the deamination-dependent mode for L1 restrictions. These ambiguities arose because no G→A mutations, which indicate the occurrence of C→U deamination on the L1(-) cDNA, could be detected on the coding strand (i.e., (+) cDNA) of the integrated L1. However, none of these cell culture studies blocked the action of UNG [40, 325]. Therefore, the unsuccessful attempts to uncover A3-induced mutations were likely due to the degradation of uracil-containing cDNA intermediates by the UNG and APE-mediated cellular repair machinery [103].

We reasoned the A3-induced deaminations could contribute to L1 restriction, as traces of strand-specific G-to-A mutational signature have been identified in several endogenous retroviruses [329, 422, 423] and the sequences of mammalian L1s by genome-wide DNA editing analysis [314]. To address this question, we used a cell culture-based L1 retrotransposition assay to assess human L1's restriction abilities by A3 proteins [424, 425]. The results presented here focused on three A3 family members, A3A, A3G and A3H, as it has been shown that A3A is able to restrict human L1 retrotransposition whereas A3G is unable to restriction L1 efficiently [326, 334, 390, 391]. There are up to at least seven A3H haplotypes exist in humans and haplotype II (Hap II) and haplotype V (Hap V) comprise the majority of the stable haplotypes (see Chapter 7). Although previous research demonstrated stable A3H could inhibit L1 [45, 289], A3H polymorphism confounded the analysis and no studies have clarified which stable A3H haplotypes contributed to L1 restrictions. In this study, we have demonstrated
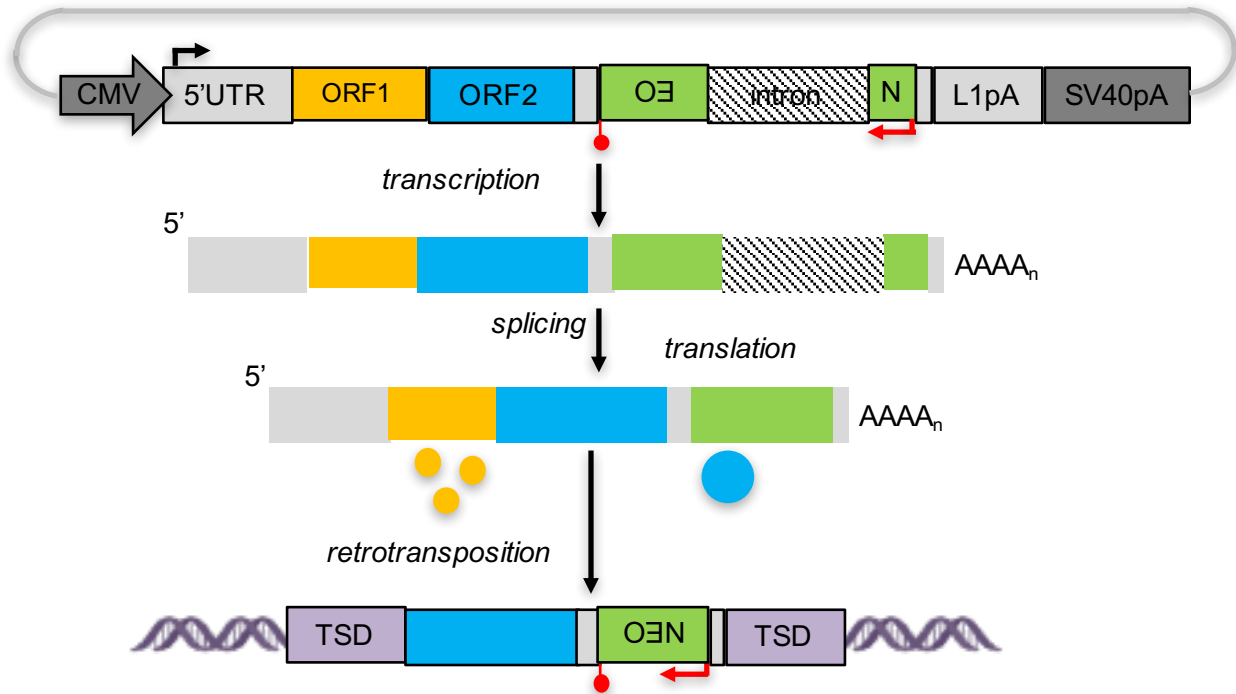
Figure 9.1 Schematic for L1 retrotransposition assay. A full-length retrotransposition-competent L1 (~6 Kb) contains a neo-reporter cassette (green box) in the 3'UTR region. The L1 transcription initiates with an internal RNA polymerase II promoter within the 5'UTR (black arrow) or the CMV promoter from pCEP4 plasmids backbone [424]. Transcription is terminated at an SV40 polyadenylation signal. The reporter gene, neomycin phosphotransferase (backwards NEO), is in the opposite transcription orientation with respect to L1. Neo gene is interrupted by an intron (hatched box), which is in the same transcriptional orientation with L1. The neo-reporter cassette has its own heterologous promoter (upside down red arrow) and polyadenylation signal (upside down red lollipop). Once transcribed, the intron is spliced out from the precursor mRNA. The resulting bi-cistronic mRNA is translated, generating two proteins, ORF1 proteins (ORF1p, orange circles) and ORF2p (blue circle). ORF1p and ORF2p preferentially act in cis and bind to their encoding mRNA, giving rise to the L1 ribonuclear protein (RNP) complex, an important L1 retrotransposition intermediate. The RNP complex then shuttles to nucleus for L1 TPRT and integration event[426, 427]. Only upon reverse transcription and integration into a genomic locus can the neo gene be expressed to confer resistance to G418. The retrotransposed L1 is usually 5' truncated and integration sites are usually flanked by target site duplications (TSD) sequences.

A3A inhibits L1 replication using a deamination-dependent mechanism and the recovered L1 retrotransposition events from Hela cells co-expressing A3A and the UGI protein contained A3A-mediated G-to-A deamination signatures. Both A3H Hap II and Hap V suppress L1 to a similar level but the inhibition of L1 did not involve DNA deamination.

9.3 Results

9.3.1 The effects of A3s on L1 retrotranspositions in mammalian cultured cells. To study the influence of A3 proteins on L1 retrotranspositions, a pre-established cell based retrotransposition assay was performed [424, 425] (Fig. 9.1). To detect rare retrotransposition events, the 3' untranslated regions (3'UTRs) of a human L1 retrotransposon vector (pJM101/L1.3) is marked with a *neomycin phosphotransferase (neo)* reporter cassette, which, when expressed, can convey resistance to G418, a neomycin analog that is used in mammalian cultured cells. The *neo* reporter cassette has its own promoter and polyadenyation signal, and it consists of an antisense *neo* gene disrupted by an intron of the *γ-globin* gene, which is in the same transcriptional orientation as the L1 element (Fig. 9.1). Because the *neo* gene is rendered inactive by the presence of an intron, and G418-resistant (G418^R) cells will only arise when a transcript initiated from the L1 promoter is spiced, reverse transcribed, its cDNA is re-inserted into the host chromosome, and the *neo* gene is transcribed from its own CMV promoter. The L1 assay relies on *neo* expression as a readout and the rise of a G418^R colony demonstrates a successful L1 retrotransposition event. Hela cells were utilized in this assay, as these cells contain a low endogenous L1 copy number (~25 copies per cell) [332], can accommodate high levels of L1 retrotransposition events from ectopically expressed L1 reporter plasmids [424, 428] and do not express endogenous A3A, A3G and A3H proteins [332], thus providing an ideal system for measuring the direct effects of ectopically expressed A3 proteins on L1 retranspositions.

In order to study the effects of A3 proteins on L1 retrotransposition, Hela cells were co-transfected with L1 reporter plasmid pJM101/L1.3 in the presence and absence of plasmids expressing V5-tagged A3s. Three days post transfection, the cells were subjected to G418 selection for 11 days and the G418^R colonies were stained and counted. Transfection of L1 vector alone with no A3 expression generates a quantifiable amount (50-150 colonies per well in a 6-well dish) of G418^R colonies (Fig. 9.2A, no A3). In the presence of A3A and A3G, L1
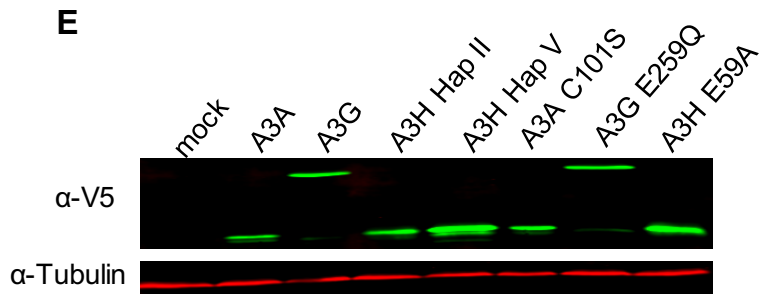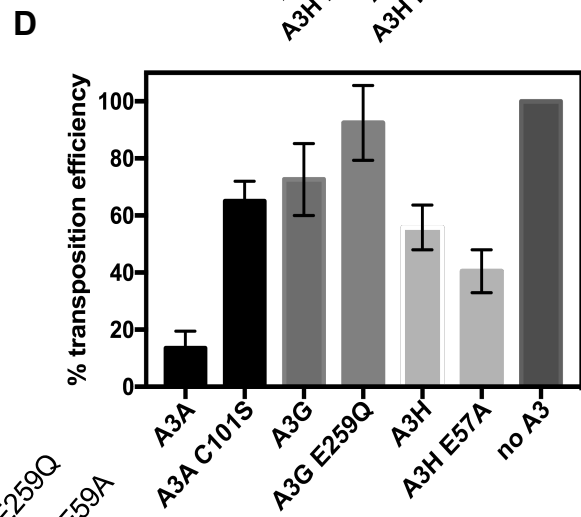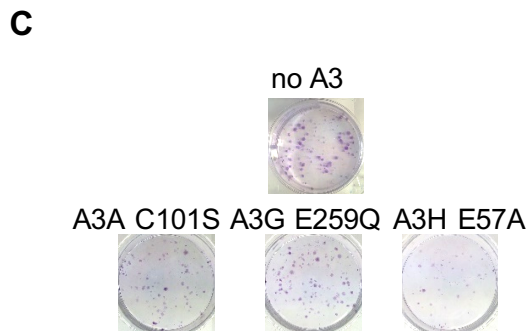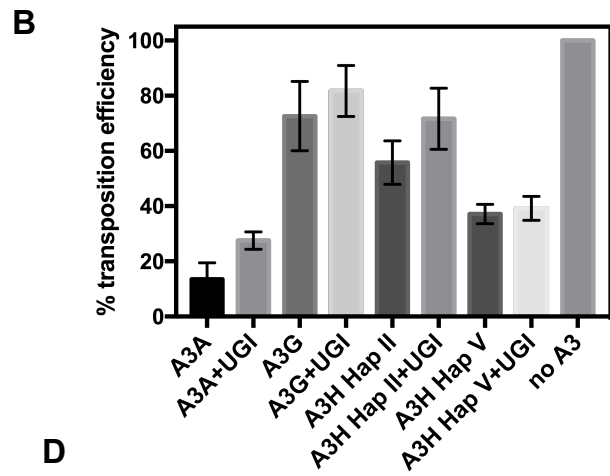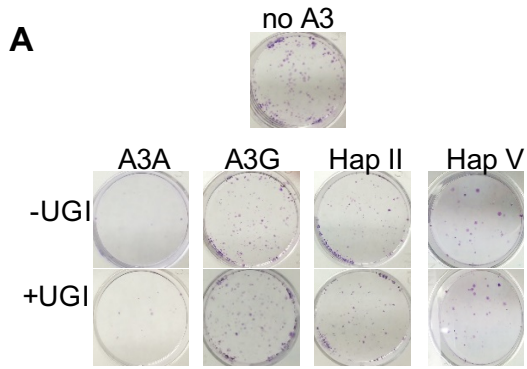
Figure 9.2 Relative L1 retrotransposition frequencies in the presence of wild type and catalytic mutants of A3A, A3G and A3H in cultured Hela cells. (A) Representative experimental results of the neo-based L1 retrotransposition assay in the presence of wild type A3 proteins. Hela cells were co-transfected with 1μg of both a neo^R-based L1 reporter plasmid (pJM101/L1.3) and the expression plasmids encoding A3A, A3G, A3H (Hap II and Hap V) or an empty vector (pcDNA 6A). Transfections were performed in the absence or presence of 1 μg of an expression plasmid for uracil DNA glycosylase inhibitor (UGI). Cells were subjected to G418 selection and resistant colonies were fixed and counted 14 days post-transfection. (B) Quantification of the effects of various A3s on L1 transposition efficiencies in the presence or absence of UGI as shown in (A). The relative retrotransposition efficiencies were calculated by counting the fixed and stained G418^R colonies and setting the value determined from control cells co-transfected with a L1 vector and an empty pcDNA 6A vector at 100%. Date were derived from at least three independent transfection experiments and are represented as the means ± standard deviations. (C) Representative experimental results of the neo^R-based L1 retrotransposition assay in the presence of A3 catalytic mutants and A3 catalytic mutants (A3A C101S, A3G E259Q, and A3H E59A) demonstrate varying effects on L1 retrotransposition. (D) Quantification of the effects of various A3 catalytic mutants on L1 transposition efficiencies as shown in (A). The relative retrotransposition efficiencies of the A3 catalytic mutants were calculated by counting the fixed and stained G418^R colonies and setting the value determined from control cells co-transfected with a L1 vector and an empty vector at 100%. Date were derived from at least three independent experiments and are represented as the means ± standard deviations. (E) The cellular expression of V5-tagged A3 proteins in Hela cells was confirmed by immunoblotting using anti-V5 antibodies. Each lane was loaded with 40μg cell lysates extracted from Hela cells 72-hrs post co-transfection of A3 expression plasmid and pJM101/L1.3. α-Tubulin (~50 kDa) expression was used as a loading control.
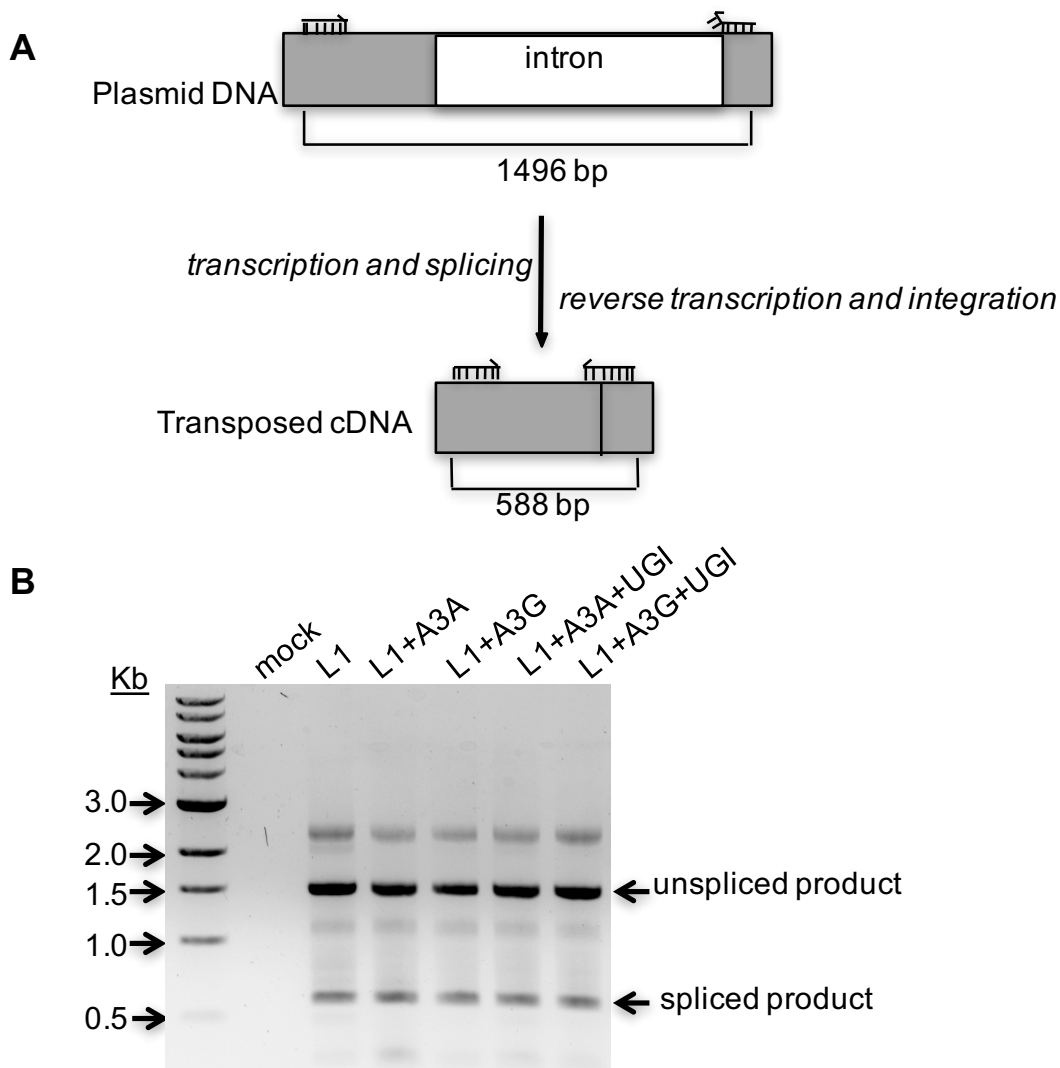
Figure 9.3 Rationale of analyzing A3-mediated deaminations footprints in L1 DNA by examining transposed neo gene. (A) Schematic of the neo-cassette before and after L1 retrotransposition with specific primers and the size of the expected PCR fragments indicated. The splice-specific reverse primers were designed to span the intron by annealing to exonic sequence 5 bp upstream and 16 bp downstream. (B) A representative ethidium bromide-stained 1% agarose electrophorses gel of the PCR products amplified using the primers shown in (A), with the 1496 bp fragment specific for the non-transposed parental copies (unspliced product) and the 588 bp fragment specific for the retrotranscribed spliced copies (spliced product) indicated. A spurious band with the size of ~2.5 kb was occasionally observed depending on the transfection conditions.

L1 retrotransposition efficiency decreased to 13.5% and 72.6%, respectively (Fig 9.2B), suggesting A3A potently inhibits L1 retrotransposition in cultured cells and A3G does not appear to have a significant inhibitory effect on L1, in agreement with previous studies [326, 332, 334, 391]. A3H Hap II and Hap V decreased the L1 retrotransposition efficiency to 55.8% and 37%, respectively (Fig 9.2B), suggesting two A3H haplotypes both impose a moderate effect on L1 mobilizations. If A3-induced L1 inhibition involves deamination activity, then the observed decrease in L1 retrotransposition efficiency was due to a combination of the L1 intermediates degradation by host DNA repair enzymes and the functional inactivation of L1 through missense or nonsense mutations induced by A3s. As a result, we first investigated whether the inhibition of L1 by A3s involved deamination by creating a number of catalytic mutants that have been reported to have abolished deaminase activities (Fig 9.2C, A3A C101S, A3G E259Q, A3H E57A) [57, 213, 334]. These catalytic mutants were created by mutating some key amino acids in A3s, such as C101 in A3A that is involved in coordinating active site Zn^{2+} ion, or E259 in A3G and E57 in A3H that act as a proton shuttle and participate directly in catalysis [430, 431]. Of note, we undertook this investigation by only studying the catalytic mutant of Hap II. Because two A3H haplotypes demonstrated similar L1 restriction abilities in the retrotransposition assay (Fig. 9.1), presumably mutating E57 to an alanine in Hap II and Hap V would lead to the same result. The inhibition of L1 imposed by A3A seemed to require its deaminase activity, as a 4.6-fold increase in retrotransposition efficiency was observed when comparing that of wild type A3A to A3A C101S (Fig 9.2D, retrotransposition efficiency increased from 13.5% to 62.2%). The L1 retrotransposition efficiency in the presence of A3G E259Q increased to 98.8% (Fig 9.2D) (Fig 9.2D, retrotransposition efficiencies increased from 72.6% to 98.8%). No significant increase was observed when compare to wild type A3G, further confirming A3G only imposes a slight effect on L1. Mutating E57 to an Ala made no significant difference to the L1-restriction abilities of A3H (Fig 9.2D, 40% for A3H E57A; 55.8% for Hap II and 37% for Hap V). This finding provides evidence that although two haplotypes both impose moderate inhibitory effects on L1, deaminase activity is not required for A3H-mediated L1 restriction. To exclude the possibility that the observed differences in retrotransposition efficiencies was due to different cellular A3 level, we assessed A3s (both wild type and catalytic mutants) expression level by Western blot in a parallel set of co-transfected Hela cell cultures (Fig. 9.2E). Immunoblotting results of cell

extracts harvested 72 hrs after transfection showed that a significant and comparable amount of A3 proteins can be detected in Hela cells and the sizes were consistent with the size of each protein, i.e., A3A and A3H is ~22 kDa and A3G is ~46 kDa (Fig. 9.2E).

L1 inhibition imposed by A3s could be partly alleviated by introducing the uracil DNA glycosylase inhibitor (UGI) protein from bacillus phage PBS2 that functionally inhibits UDG by dissociating the DNA/UDG complex [103, 432]. To further verify whether the inhibition of L1 retrotransposition by A3s is deamination-dependent, UGI expression vector was co-transfected along with L1 and A3 plasmids and the retrotransposition efficiencies were examined (Fig. 9.2, A and B). The inhibition in L1 retrotransposition efficiency by A3A was alleviated 2-fold when UGI was expressed (Fig. 9.2B, from 13.5% to 27.5%), indicating the L1 inhibition imposed by A3A indeed involves deaminations. UGI expression did not cause a significant increase in L1 retrotransposition efficiencies for A3G and A3H (Fig. 9.2B. For A3G, the retrotransposition efficiencies in the presence and absence of UGI are 72.6% and 82.8%; for A3H Hap II, the retrotransposition efficiencies in the presence and absence of UGI are 55.8% and 71.6%; for A3H Hap V, the retrotransposition efficiencies in the presence and absence of UGI are 37% and 39.2%), indicating the slight decrease in retrotransposition efficiency in the presence of A3G expression (Fig. 9.2B, ~30%) was not due to its deaminase activities and for A3H, the moderate inhibitory effects on L1 mobilizations was not caused by A3H-induced deaminations.

9.3.2 A3-catalyzed mutation analysis in transposed neo genes. Next, we wish to examine any potential A3-mediated deamination footprints by sequencing transposed L1. Because the detection of L1 retrotransposition event in the L1 assay relies on the expression of the spliced *neo* gene and the neo-cassette located within the 3'UTR remains single-stranded for the longest time since it is furthest from the priming site for (+)DNA, clones that contain a heavily mutated *neomycin phosphotransferase* may fail to convey G418-resistance and will be lost during the antibiotic selection process. This leads to an underestimation of A3 enzymes' deamination potency. To circumvent this problem, instead of sequencing the L1 region of each individual G418[R] clone, cellular DNA was extracted and used to for PCR reactions three days post-transfection prior to G418 selections. Because an intron was inserted into the *neo* gene and a spliced *neo* gene could be produced only after an L1 integration event (Fig. 9.1), we designed primers that anneal to exon regions spanning the intron, enabling amplification of PCR products with different lengths that correspond to the spliced and unspliced products (Fig. 9.3A). UGI was

introduced to these experiments to promote the recovery of mutations. Several bands were revealed in the PCR amplification products even with the primers that intend to eliminate the amplification of intron sequences (Fig. 9.3B), in agreement with the previous observation that retrotransposotion is a rare event [334]. Sequencing confirmed the 588 bp species corresponded to the spliced, transposed *neo* products, whereas the 1496 bp species corresponded to the PCR amplification of the unspliced, full-length *neo* (data not shown).

Sequencing the transposed intronless *neo* genes suggest in the absence of A3, the background G-to-A mutation rate is 0.24 mutations/kb (Fig. 9.4A, no A3). In the presence of A3A, the G-to-A mutation rate increased to 0.68 mutations/ kb (Fig. 9.4B), this 3-fold increase in mutation frequency partially explains why A3A is very potent at inhibiting L1 (Fig. 9.2B) and despite the degradation of uracil-containing L1 intermediates, some L1 intermediates can still escape the degradation of DNA repair enzymes and integrate into the host chromosome. Remarkably, when UGI was co-transfected along with L1 and A3 vectors, a 5.6-fold increase in G-to-A mutation rate was observed (Fig. 9.4C, A3A+UGI, the G-to-A mutation rate is 3.8 mutations/ kb). This suggests introducing UGI blocked UNG and allowed more G-to-A mutations to be uncovered. Very few C-to-T mutations were observed on the transposed L1 coding strands, excluding the possibility of A3A deaminating L1 RNA (Fig. 9.4, B and C). Because A3A prefers to mutate G-to-A within a 5'GA context (with the underlined G being mutated) [206], we next investigated whether these G-to-A mutations have occurred within A3A's preferred deamination motif. In the presence of A3A, out of the 16 G-to-A mutations we recovered in *neo* gene (Fig. 9.4B), only two mutations are within the 5'GA to 5'AA context, and mutations in the 5'GC to 5'AC, 5' GG to 5'AG, 5'GT to 5'AT motifs were observed (data not shown). Interestingly, in the presence of UGI, the majority of the G-to-A mutations have occurred within the 5'GA to 5'AA context (Fig. 9.4D), suggesting these mutations are related to A3A-induced deaminations. Consistent with the previous observation that A3G has little inhibitory effect on L1 activity (Fig. 9.2, B and D), a G-to-A mutation frequency of 0.16 mutations/kb similar to that of the background was observed in *neo* gene (compare Fig 9.4A and 9.4E, G-to-A mutation frequency of 0.24 mutations/kb *v.s* 0.16 mutations/kb). UGI expression did not cause an increase in G-to-A mutation frequency for A3G (Fig. 9.4F, G-to-A mutation frequency of 0.1 mutations/kb), suggesting the slight inhibitory effect of L1 retrotransposition by A3G (Fig. 9.2B, 72.6%) we observed was not caused by A3G-mediated deamination. Taken

A   **No A3**

To

	A	T	C	G
A		0	17	0
T	0		0	0
C	0	0		0
G	**4**	0	0	

From

n=16,464 bp

B   **A3A**

To

	A	T	C	G
A		2	33	8
T	1		3	3
C	3	2		11
G	**16**	1	13	

From

n=23,520 bp

C   **A3A+UGI**

To

	A	T	C	G
A		0	33	2
T	3		0	0
C	1	1		2
G	**49**	0	0	

From

n=12,804 bp

D

```
H11   AAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAG
E10   AAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAG
E2    AAGGTAAGATAACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAG
F8    AAGGTAAGATAACAGGAAATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAG
G12   AAGGTGAGATGACAGAAAATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAG
Ref   AAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAG
      ***** **** **** * *********************************
```

E   **A3G**

To

	A	T	C	G
A		2	25	1
T	1		0	1
C	0	1		0
G	**3**	0	0	

From

n=19,404 bp

F   **A3G+UGI**

To

	A	T	C	G
A		0	15	0
T	0		0	0
C	0	0		0
G	**2**	0	0	

From

n=19,992 bp

Figure 9.4 A3A, but not A3G, promotes G-to-A mutations in the transposed neo genes. (A) to (C) and (E) to (F): nucleotide preference substitutions detected in the transposed neo genes in the presence of A3A and A3G and UGI. n, total number of nucleotides analyzed. The number of G-to-A transition mutations in the (+) strand cDNA sequence (indicative of C-to-U deamination events on (-) strand cDNA) is highlighted. (A) L1 alone (with no A3 transfected) demonstrated a background G-to-A mutation frequency of 0.24 mutations/kb. (B) In the presence of A3A, the G-to-A mutation frequency is 0.68 mutations/kb. (C) UGI expression uncovered the evidence of A3A-mediated deamination events in transposed *neo* gene. In the presence of A3A and UGI, the G-to-A mutation frequency is 3.8 mutations/kb. (D) DNA sequencing data from a subset of clones showing the sequence context of G-to-A mutations induced by A3A in the transposed *neo* gene in the presence of UGI. Asterisks denote homology. (E) A3G demonstrated the G-to-A mutation frequency is 0.16 mutations/kb. (F) UGI expression did not allow more A3G-mediated deamination events in transposed *neo* gene to be revealed. In the presence of A3G and UGI, the G-to-A mutation frequency is 0.1 mutations/kb.

A A3H Hap II

From \ To	A	T	C	G
A		0	20	0
T	1		0	0
C	0	0		0
G	**3**	0	1	

n=17,052 bp

B A3H Hap II +UGI

From \ To	A	T	C	G
A		0	30	0
T	0		0	0
C	0	1		3
G	**3**	0	2	

n=22,932 bp

C A3H Hap V

From \ To	A	T	C	G
A		0	32	4
T	1		0	2
C	2	0		1
G	**3**	0	4	

n=22,932 bp

D A3H Hap V +UGI

From \ To	A	T	C	G
A		1	38	3
T	0		0	1
C	1	0		4
G	**4**	1	2	

n=24,108 bp
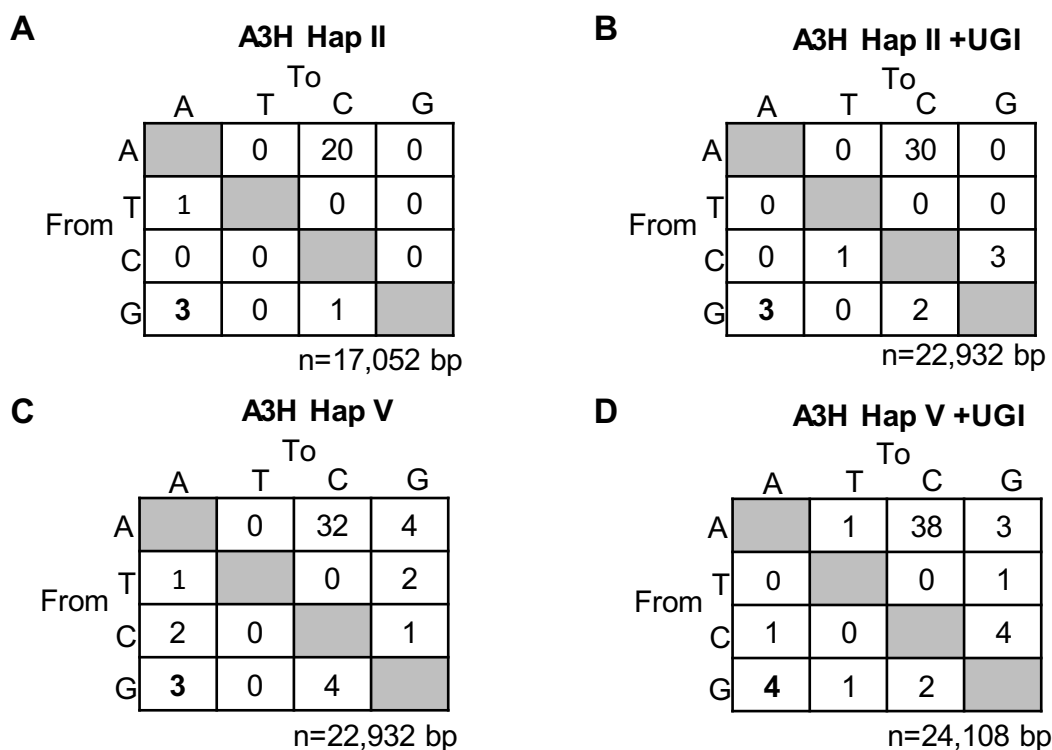
Figure 9.5 A3H Hap II and Hap V do not appear to promote any significant numbers of G-to-A mutations in the transposed neo genes. (A) to (D): nucleotide preference substitutions detected in the transposed neo genes in the presence or absence of A3H Hap II, Hap V and UGI. n, total number of nucleotides analyzed. The number of G-to-A transition mutations in the (+) strand cDNA sequence (indicative of C-to-U deamination events on (-) strand cDNA) is highlighted. (A) In the presence of A3H Hap II, the G-to-A mutation frequency of 0.18 mutations/kb. (B) UGI expression did not uncovered any more A3H Hap II-induced deamination events in transposed *neo* gene. In the presence of A3H Hap II and UGI, the G-to-A mutation frequency is 0.13 mutations/kb. (C) In the presence of A3H Hap V, the G-to-A mutation frequency is 0.13 mutations/kb. (D) UGI expression did not uncovered any more A3H Hap V-induced deamination events in transposed *neo* gene. In the presence of A3H Hap V and UGI, the G-to-A mutation frequency is 0.16 mutations/kb.

together, these data strengthen the findings that A3A uniquely inhibits L1 retrotransposition using a deamination-dependent mode, whereas A3G does not seem to affect L1 mobilization. Because A3H catalytic mutant E57A restricts L1 as potently as the wild type A3H (Fig. 9.2D) suggesting A3H does not require deamination activity to restrict L1, we hypothesized no G-to-A mutations should be uncovered on transposed *neo* gene in the presence of A3H, in contrary to what has been observed for A3A (Fig. 9.4, B and C). Indeed, no enhanced rates of G-to-A mutations have been detected in *neo* genes in the presence of A3H Hap II or Hap V (Fig. 9.4A, background G-to-A mutation rate is 0.24 mutations/kb; Fig 9.5A, Hap II, G-to-A mutation frequency is 0.18 mutations/kb; Fig 9.5B, Hap V, G-to-A mutations frequency is 0.13 mutations/kb), confirming A3H does not utilize a deamination-dependent mechanism to inhibit L1. In addition, introducing UGI did not increase the number of G-to-A mutations for both Hap II and Hap V (Fig. 9.5, B and D, in the presence of UGI, A3H Hap II induced 2 G-to-A mutations whereas A3H Hap V induced 3 G-to-A mutations). None of the G-to-A mutations appeared within a 5'GA to 5'AA context (data not shown). Neither haplotypes appeared to be able to deaminate L1 RNA, as no apparent C-to-T mutations with the signature context were found (Fig. 9.5).

9.4 Discussion and future directions

In this study, we showed different A3 family members, i.e., A3A, A3G and A3H, have varying abilities to inhibit the replication of retrotransposon L1 in a cultured cell assay. Specifically, A3A acts as a potent inhibitor for L1 retrotransposition, whereas A3G does not appear to inhibit L1 (Fig. 9.2, A and B), in accordance with previous studies [40, 325, 326, 332, 334, 390]. Previous studies showed the stabilities of A3H haplotypes affect their abilities to inhibit L1. Specifically, the ΔN15 and G105-bearing stable haplotype(s) have a strong ability to restrict L1 [45, 289], whereas N15 and R105-bearing unstable haplotype have lost this function [334, 332]. This knowledge needed to be revisited as our understanding of A3H polymorphisms deepened. We found that two prevailing stable A3H haplotypes, Hap II and Hap V, both demonstrated a moderate inhibitory effect on L1 mobilizations. This is the first time, to the best of our knowledge, that the effects of stable A3H haplotypes on L1 have been comparatively examined. Numerous studies have reported the inhibitory effects of different A3 family members on L1, and in general a deamination-independent L1 restriction model has been proposed, as no

G→A mutations related to A3-related deamination events have been detected with L1 restriction, even by A3A [325, 326, 332-334, 390, 391, 421]. This might be largely due to the edited L1s being degraded by cellular DNA repair enzymes. Indeed, Richardson *et al*., studied the effects of A3A on L1 retrotransposition in a cell line that stably expresses UGI, and this allowed the recovery of A3A-mediated deamination events in the retrotransposed L1 sequences [337]. Inspired by this study, we set out to investigate the mechanisms different A3s utilized to restrict L1s by conducting experiments in the presence of UGI that has been introduced in a transfected plasmid.

A3A has been demonstrated to be the most potent inhibitor of L1 mobilizations where it can restrict L1 retrotransposition frequencies by 75% to 95%, depending on the transfection conditions [326, 332, 334, 337, 391]. In our system A3A inhibits L1 retrotransposition by 85% (Fig. 9.2D) and deamination-deficient mutant, A3A C101S, only moderately restricts L1 mobility (Fig. 9.2D, A3A C101S, L1 retrotransposition efficiency decreased by 40%). These findings indicate that the deaminase activity of A3A contribute greatly to L1 inhibition. It should be note that A3A has been reported to convey cytotoxicity on cells [433-435] and this could be responsible for the significant reduction in G418^R colonies we observed in the L1 assay in the presence of A3A expression. A3A-mediated toxicity on cells likely stems from its ability to gain access to ss genomic DNAs that are transiently exposed in transcription bubbles and replication forks, leading to cell cycle perturbation [435]. This concern could be addressed by transfecting cells with A3A proteins expressed in a plasmid that carries G418-resistance (such as pcDNA3.1 from Invitrogen) to rule out nonspecific cytotoxicity A3A has contributed to L1 inhibition. To test whether excision of uracil bases and APE-mediated L1 cDNA degradation contributed to A3A-induced L1 inhibition, we performed L1 assay in the presence of a bacteriophage-derived UGI that inhibits UNG activity [103] and found that UGI expression alleviated A3A-mediated L1 inhibition by 2-fold (Fig. 9.2B, from 13.5% to 27.5%), consistent with what Richardson *et al*. have reported using transient transfection [337]. The alleviation of inhibition by UGI is inefficient with transient transfection, as A3A still strongly inhibits L1 in the presence of UGI by ~70% (Fig. 9.2B), and there are a number of explanations for this observation. First and foremost, the inhibition of UNG activity by UGI was likely not 100% due to transient transfection method. To make sure UNG activity is fully inhibited before the co-transfection of L1 and A3 vectors, a cell line that stably express UGI could be adopted, as it has been suggested that with a cell line

that stably express UGI, A3A almost lost its ability to restrict L1 and only 20% L1 inhibition was observed [337]. Alternatively, UGI expression vector could be transfected one day before the co-transfection of UGI, L1 and A3 vectors to boost the expression of UGI. Secondly, mammalian cells express two nuclear UNGs with activities on ssDNA: UNG2, which is inhibited by UGI, and SMUG1, which is not affected by UGI [436-438]. Hence, SMUG1 could contribute to degrading the lesion-containing L1 substrates when UNG2 is blocked by UGI. Nevertheless, to further prove UNG is involved in degrading ss L1 intermediates at which A3A has been acted upon, one could test the inhibitory effects of L1 by A3A C101S in the presence of UGI and UGI expression should have little effect on L1 activity in assays conducted with A3A C101S.

Customarily, to verify whether the mode of L1 inhibition involves A3-mediated deaminations, the integrated genomic DNA present in individual G418^R clones is sequenced to examine for the presence of G-to-A mutations on L1 coding strand, which serves as a direct evidence for A3-mediated deamination events [40, 337]. This approach tends to produce an underestimated low G-to-A mutation frequency, because clones that have heavily mutated genomes, which rendered *neo* gene to be non-functional, would be lost during the G418 selection process. We chose to PCR and sequence the intronless *neo* DNA that can only be generated in a spliced and reverse-transcribed L1, before G418 selection, to preserve the deamination events that have generated high numbers of mutations (Fig. 9.3). Interestingly, we noticed a high percentage (81%) of A→C transversion mutations present in the background (Fig. 9.4A, no A3) and G→A transition mutations account for the second most frequent mutation types occurred (19%) (Fig. 9.4A, no A3). A large amount of A→C mutations were persistently observed across different transfection conditions (Fig. 9.4, A-C, E-F and Fig. 9.5), suggesting their occurrence is not due to a specific A3 enzyme and may be an insertion bias of the L1 reverse transcriptase. A3A appeared to be a strong inhibitor of L1, as A3A induced a G-to-A mutation frequency that is 3 times above the background G-to-A mutation level (compare Fig. 9.4A and Fig. 9.4B, G-to-A mutation frequency of 0.24 mutations/kb *v.s* 0.68 mutations/kb). In contrast, an A3 that is not competent at inhibiting L1, such as A3G, induces a G-to-A mutation frequency that is similar to that of the background (compare Fig. 9.4A and Fig. 9.4E, G-to-A mutation frequency of 0.24 mutations/kb *vs.* 0.16 mutations/kb). Due to the limited numbers of G-to-A mutations revealed in the presence of A3G (Fig. 9.4, E and F), no dinucleotide sequence specificity can be examined. Remarkably, the G→A mutation profile induced in the presence of UGI and A3A demonstrated a

16-fold increase in the G→A mutations rate as compared to that of the background (compare Fig. 9.4A and Fig. 9.4C, G-to-A mutation frequency of 0.24 mutations/kb *vs*. 3.8 mutations/kb), suggesting introducing UGI blocked UNG activity and allowed a significant amount of C→U lesion-containing L1 cDNA to escape degradations by cellular repair enzymes. It is noteworthy that in the presence of A3A and UGI, a large number of G→A mutations occurred in the 5'GA to 5'AA context, consistent with the preference of A3A to deaminate 5'TC motifs in (-)DNA. Few deamination hotspots were visible after inspecting the *neo* sequences (Fig. 9.4D), consistent with A3A being non-processive [206].

Taken together, we propose a model for how A3A is involved in the L1 life cycle and inhibits L1 retrotransposition. During TPRT, the transiently exposed (-) strand L1 cDNA becomes the substrates for A3A-catalyzed deamination. Usually, if not repaired promptly the uracil-containing L1 cDNA intermediates would be vulnerable to the combined actions of cellular DNA repair enzymes, such as UNG and APE, and this leads to either the complete degradation of L1 cDNA or the integration of 5' truncated L1 cDNA that lacks any evidence of A3A-mediated deaminations. By introducing an ectopically expressed UGI, A3A-mediated deamination events become uncovered in retrotransposed L1 sequences. The significant decrease of L1 retrotransposition efficiency in the presence of A3A (Fig. 9.2B) was attributed by the combined action UNG/APE-mediated L1 cDNA degradation and the mutations induced by A3A. The former may contribute more to the general inhibitory effects observed. One way to clarify the contribution of these two inhibitory actions is to introduce UGI, thus blocking the L1 degradation pathway and only allows the effects of A3A deamination to be revealed. In our system, blocking UGI only alleviated the inhibition by 2-fold and A3A still strongly inhibits L1 activity up to 27.5% (Fig. 9.2D). This neither suggests the rest of ~70% of L1 inhibition is caused by A3A-mediated deaminations or reflect the contributions UNG/APE-mediated degradation pathway made towards L1 inhibition, as the inhibition of UND by UGI was likely incomplete based on our transfection methods, as discussed above. Due to A3A catalytic mutant only restored L1 activity to ~60% (Fig. 9.2D), our model does not necessarily exclude possibilities of deamination-independent mode of actions, such as A3A inhibiting TPRT by inhibiting L1 RT or EN activity; nevertheless, this deamination-independent activity would be only 40% of the total restriction activity at best.

Previously through the characterization of A3H Hap II and Hap V we found that two A3H Haplotypes differ in their RNA binding abilities [210]. Specifically, A3H Hap II demonstrated stronger binding affinities for HIV 5' UTR and Alu RNA than A3H Hap V. These two haplotypes did not show any differences in HIV restriction capacities [210]. Hence, it seemed interesting to investigate whether the difference in RNA binding ability could contribute to a difference in the abilities for two haplotypes to restrict L1, as it has been demonstrated the ability to bind RNA is crucial for L1 inhibiting abilities to some A3 family members, such as A3C and A3D [335, 336]. We found A3H Hap II and Hap V are able to suppress L1 retrotransposition to a moderate level (Fig. 9.2, A and B) and there are no apparent differences in their L1 restriction abilities (Fig. 9.2B, Hap II restricts L1 retrotranspostion efficiency to 55.8%, Hap V restricts L1 retrotransposition efficiency to 33.7%). Three lines of evidence suggest both A3H haplotypes inhibit L1 mobilization in a manner that does not involve deaminations. Firstly, we found Hap II and V produced a similar G→A mutation profile and the genomic L1 retrotransposition event that occurred in the presence of both A3H haplotypes are devoid of any G-to-A mutations (Fig. 9.5). Secondly, UGI expression had little effect on L1 activity in assays performed in the presence of two A3H haplotypes (Fig. 9.2B). Lastly, A3H catalytic mutant E57A still decreases L1 retrotransposition efficiency (Fig. 9.2D). Altogether, these findings suggest two stable A3H haplotypes utilize one or multiple deamination-independent mechanisms to restrict L1 activities.

At early stage of L1 life cycle, L1 forms a ribonucleoprotein (RNP) complex as a retrotransposition intermediate by associating its own mRNA with ORF1p and ORF2p in the cytoplasm [426, 427]. The RNP complex then shuttles back to the nucleus and L1 replication requires reverse transcription from of its own RNA genome using reverse transcriptase encoded by ORF2p. For all A3s that have been reported to be capable of restricting L1, cellular localization of an A3 protein does not necessarily dictate its ability to restrict L1, as different mechanisms can be utilized by A3s to inhibit various stages of L1 mobilizations. Cytoplasmic A3s can inhibit L1 by binding ORF1p in an RNA-dependent manner and prevent the nuclear import of L1 RNP complex [335]. In contrast, A3s that locate in nucleus [333, 439] have been reported to inhibit L1's TPRT process either by deaminating (-)cDNA or physically impair RT's activity. Based on these observations and that both A3H Hap II and Hap V are primarily cytoplasmic proteins [45], we hypothesize A3H could inhibit L1 activity using a sequestration

mechanism similar to that of A3C [335]. As a first step to investigate this hypothesis, one could test whether A3H interacts with ORF1p by co-immunoprecipitation and whether this interaction can be disrupted by RNase treatment. The nature of RNA molecule that mediates A3H-ORF1p interaction can be investigated by immunoprecipitating A3H after transfection and detecting RNA in the eluted samples by quantitative PCR. L1 RNA and 7SL RNA serve as good candidate as 7SL RNA has been reported to be associated with L1 RNPs [440]. In addition, immunofluorescence microscopy can be used to examine whether A3H and ORF1 co-localize in the same cytoplasmic compartment, such as stress granule.

9.5 Material and Methods

9.5.1 Plasmids, cloning and site-directed mutagenesis. L1 expression plasmid pJM101/L1.3-Neo (construct enables resistance to G418) was kindly provided by Dr. John V. Moran (University of Michigan, Ann Arbor). Mutants of A3 (A3A C101S, A3G E259Q, A3H E57A) were made using the WT A3 construct for site-directed mutagenesis (QuickChange site-directed mutagenesis protocol, Stratagene). cDNA of WT A3A, A3G and A3H and mutants were cloned from pcDNA3.1 (Invitrogen) into pcDNA6-V5/His.A using XbaI and XhoI cloning sites. A3 proteins expressed from this vector have V5-His fusion tags present at the C terminus. Uracil DNA glycosylase inhibitor (phage gene synthesized by GeneScript) was subcloned into pcDNA6-V5/His. A stop codon was introduced at the end of the UGI sequence so native UGI can be expressed. All constructed plasmids were verified by sequencing.

9.5.2 Cells. Human Hela cells (ATCC® CCL2-1™) were grown in Dulbecco's modified Eagle medium (DMEM) (HyClone) containing 10% fetal bovine serum (FBS) (Gibco) and maintained in a humidified 37 °C incubator with 5% CO_2. To test the minimal amount of G418 required to kill the untransfected Hela cells, an antibiotic killing curve was performed. In detail, Hela cells (2×10^5 cells/well) were seeded into a 6-well plate 20 hrs before adding G418. The next day, a titrating amount of G418 (0-1000 µg/mL) was added to the media and cells were maintained for 11 days in G418-containing media. Selection media were changed every 3 days. It was determined that 400 µg/mL of G418 provides the lowest antibiotic concentration at which cell death can visualized after 2 weeks and should be used for the retrotransposition assay (see below).

9.5.3 Retrotransposition assay. Hela cells were seeded in each well of a 6-well tissue culture plate at a density of 2×10^5 cells/well in growth media. Approximately 22 hrs after plating, Hela

178

cells were co-transfected with 1μg of L1 retrotransposon encoding plasmid (pJM101/L1.3) and with 1μg of the respective A3 protein encoding plasmid. Empty vectors were used to equalize transfected amount. Transfections were performed using GeneJuice (EMD Millipore) according to the manufacturer's protocol. Media was replaced the following day. Approximately 48 h post-transfection, cells were re-seeded to 6-well plate (at 1 in 5 dilution) and the next day attached cells were subjected to selection with 400 μg/mL G418 (Life Technologies). G418 selection was carried out for 11 days and the selection media were replaced every other day. G418^R colonies were washed with 1x phosphate buffered saline (PBS) for two times, fixed with 2% paraformaldehyde/0.4% glutaraldehyde (Sigma), and stained with 0.1% crystal violet solution.

9.5.4 Immunoblotting. To detect A3 protein expression levels in L1 assay, the same transfection conditions were used as described in the retrotransposition assay. Because all A3 protein expression plasmid were constructed with a C-terminal V5 tag, the proteins could be detected using antibodies specific for V5 epitope. Approximately 72hr post-transfection, cells were washed in 1x PBS and harvested in Laemmli sample buffer (58 mM Tris (pH 6.8), 5% glycerol, 2% sodium dodecyl sulphate (SDS), 1.5% dithiothreitol (DTT)). The amount of proteins in the cell lysates was determined using Lowry assay (Total Protein Kit, Mirco Lowry, Peterson's Modification, Sigma-Aldrich). Samples were heated at 95°C for 10 min prior to SDS-PAGE, and equal amount of lysates (40 μg) were loaded. Proteins were transferred to nitrocellulose membranes using electrophoresis (100 V, 1h at 4°C) and A3 proteins in the cell lysates were probed with primary antibodies (mouse monoclonal V5, 1:1000, Sigma). Monoclonal α-tubulin antibodies (rabbit monoclonal α-tubulin, 1:1000, Sigma) were used to examine equal amount of loadings in each well. After incubation with the secondary antibodies (A3 proteins: IRDye 800-labeled goat anti-mouse (Mandel Scientific, catalog # 926-32210); α-tubulin: IRDye 680-labeled goat anti-rabbit (Mandel Scientific. Catalog # 926-68071)). Tagged proteins and loading controls were detected simultaneously on the same immunoblot by using the Licor/Odyssey system.

9.5.5 Nucleic acid extraction and PCR. Hela cells were co-transfected with 1 μg of A3 and L1 plasmids in the presence or absence of 1 μg UGI vector as described above. Cellular DNA (both genomic and episomal) from the transfected cells was extracted 72 hr post-transfection using DNAzol reagent (Thermo Fisher). DNA was treated with DpnI (New England Biolabs) for 1hr at 37°C to remove residual episomal DNA from transfected plasmids. Extracted DNA was used as

a template for PCR amplification of the spliced *neo* gene using the forward primer (G418Neoseq For) 5'-TCGGGAGCGGCGATACCGT-3' and the reverse primer (G418Neoseq Rev) 5'-CGGTGCCCTGAATGAGCTTCA-3'. The PCR reaction (50 μL) contains 0.6 μg of cellular DNA, 1x Q5 reaction buffer and 1 unit of Q5 High-Fidelity DNA polymerase (NEB), 200 μM of deoxynucleoside triphosphates (dNTPs) mixture, 0.5 μM of each primer, with an initial denaturation step at 98 °C (3 min) followed by 35 cycles of amplification (30 sec at 98°C, 30 sec at 65.5 °C and 18 sec at 72 °C). PCR products were electrophoresed on 1% agarose gel in the presence of 0.1 μg/mL of ethidium bromide and exposed using ChemiDoc imager (Bio-rad). To examine the editing potency of A3 enzymes on retrotransposed L1 without prior G418 selection, the spliced PCR product was gel purified with the GenElute Gel Extraction kit (Sigma), and cloned into pJET1.2/blunt vector (CloneJet PCR cloning kit, Thermo Scientific). After transformation into *Escherichia coli* DH5α cells, clones were picked and sent for sequencing using a kit-specific primer (pJET1.2 Forward) at the National Research Council DNA sequencing Facility (Saskatoon, SK). Sequence alignments were performed using Clustal Omega.

10.0 Discussion and Future Directions

In this thesis, I have addressed the relationship of the biochemical properties of A3H and A3G as ssDNA cytosine deaminases to their biological functions as host innate immune restriction factors against exogenous retroviruses HIV and retrotransposon L1. One major conclusion derived from our studies with the biochemical characterization of A3H in comparison to A3G is that the mechanisms an A3 uses to scan non-target DNA in search for its deamination motif, instead of enzyme specific activities, determines its catalytic efficiency. Also, A3G and A3H (A3H, Z3-type single domain enzyme; A3G, Z2-Z1-type double domain enzyme) demonstrate very different properties as ssDNA cytidine deaminases, and this can be reflected in differences in nucleic acid binding, enzyme specific activity, ssDNA scanning motion and mode of oligomerization. These biochemical differences may provide clues as to why A3G and A3H have evolved to work more efficiently in certain biological environments, but not others. This may also be part of the reason that Vif has evolved a different interaction interface with A3G and A3H, for these two enzymes use different amino acids to affect oligomerization and processivity. By studying the restriction abilities of A3G and A3H under the scenarios of HIV and L1 replication, we found that A3G and A3H are processive enzymes and can be considered equally potent at inhibiting HIV replication [45, 210]; however, in our experimental system A3G is not able to inhibit L1 mobilization whereas A3H can. Both A3G and A3H are cytoplasmic proteins [45, 96] and this has likely limited their scope of L1 inhibition to the cytoplasm. Although the specific mechanism awaits elucidation, that A3H is able to bind RNA 10-fold more tightly than A3G (Table 7.3) [140, 210] may allow A3H to interact with L1 RNA more efficiently and thus prevent L1 RNP from entering the nucleus.

A large portion of this thesis was devoted to the exploration of the ssDNA scanning mechanisms of A3G and A3H (i.e., the mechanisms of facilitated diffusion). Facilitated diffusion allows the enzymes to search for their specific targets faster than the simple diffusion-limited rate and the processivity assays we performed under the multiple turnover conditions provide a powerful tool to quantify and compare the ability of different A3s to diffuse along DNA and to locate a specific site for deamination. Understanding how A3 enzymes scan ssDNA in search for their deamination targets has been a subject that is near and dear to our heart for two reasons. First of all, as compared to the dsDNA scanning enzymes, such as transcription factors [441], DNA repair, recombination and replication proteins [442, 443] and restriction enzymes [444-

181

446], there is a paucity of mechanistic studies on ssDNA scanning enzymes. Secondly, knowledge gained on the ssDNA scanning mechanisms of A3 enzymes could be extrapolated more broadly to other DNA binding proteins with positively charged DNA binding sites and allows us to rationalize how these proteins search the genome during their varying biological activities.

The sliding and jumping motions are two facilitated diffusion mechanisms that have been well characterized for A3G [132, 133, 139, 140]. Upon the characterization of A3H haplotypes we were surprised to find that A3H is able to use a different facilitated diffusion mechanism, namely, intersegmental transfer, to translocate over longer distances on ssDNA; however, jumping and intersegment transfer mechanisms seem to be redundant to the mutagenic efficiency of A3H [210]. This conclusion was reached because A3H haplotype II that can undergo both jumping and intersegmental transfer did not restrict HIV replication more than haplotype V, which can only efficiently jump and has diminished intersegmental transfer ability [210]. The advantage conveyed by a 3-D jumping movement of an A3 enzyme on ssDNA is that during HIV reverse transcription it provides an approach for overcoming a DNA/RNA hybrid on (-)cDNA (Fig 1.2); however, the advantage conveyed by intersegmental transfer movement of A3H during HIV replication, other than providing an alternative approach for long distance translocation, remained unclear. We have demonstrated that undergoing intersegmental transfer accelerates the rate of target searching process using a bulk biochemistry assay [210] and it would be interesting to validate this finding *in vivo* or with an *in vitro* model HIV replication system with the kinetics of reverse transcription being monitored. It is also possible that intersegment transfer motions of A3H facilitate template switching events by bringing different regions of the HIV RNA or (-)DNA together, which is a process during reverse transcription where HIV RT dissociates from its original template and reassociates with another template [122, 123], thus playing a role in increasing the recombination rate and expanding viral genome diversities. A3G is unable to undergo intersegment transfer under the same experimental condition tested with A3H [210] and has been found to be unable to affect the rate of recombination [76].

To test whether an enzyme is able to undergo intersegment transfer, we designed an assay that included increasing the concentration of enzyme and ssDNA substrate while keeping the ratio of each component the same [210]. A decrease in the apparent processivity factor and an

increase in the reaction rate as the concentration of the reaction components increases would indicate the occurrence of intersegment transfer [374]. Intersegment transfer could also be examined under a more stringent condition by keeping the concentration of the enzyme constant, but adding a titrating amount of competitor DNA to the reaction [133, 442]. Regardless of the reaction set up, it should be noted that if intersegment transfer is inherent to an enzyme, this type of translocation should have already occurred at the lowest DNA concentration, despite that the probability of an enzyme translocating to a separate DNA molecule nearby is low. Under this situation, in a mechanism that resembles jumping, A3H likely undergoes intersegment transfer by translocating to a distal site on the same DNA, although more severe structural alterations in the ssDNA, such as bending or looping, are likely involved [141, 149]. It would be interesting to investigate, either kinetically or structurally, whether a direct evidence to differentiate jumping from intersegment transfer could be obtained.

Historically, the mechanism of intersegment transfer has been defined as a protein with two DNA binding sites simultaneously bound to two DNA sites and forming an "intermediate" binding state before translocating to another DNA [145, 444]. Hence, an early assumption was that being oligomeric, or at least having two DNA binding domains per polypeptide, would allow an A3 enzyme to undergo intersegment transfer efficiently. However, as shown by the findings that A3G and A3F, A3s with two DNA binding domains per polypeptide, cannot undergo intersegment successfully [133, 140, 210], this prediction is likely incorrect. Despite the protein being overall positively charged, the CTDs of A3G and A3F are negatively charged and do not favor interaction with DNA [140, 182]; therefore, the re-association with DNA might be the rate limiting step and eventually allows an attempt by A3G or A3F to undergo intersegment transfer to be resolved as a jumping movement. This does not seem to be an issue for A3s with a single DNA binding domain per polypeptide chain, such as A3H and A3C, which are overall positively charged and likely use a groove around helix 6 to interact with DNA. A consensus for the structural requirements for A3s to undergo intersegmental transfer has not been achieved. For example, the ability of A3H to perform intersegment transfer depends on its ability to dimerize and the specific amino acid residue E178 on helix 6 [210]. The ability of A3C to perform intersegment transfer depends only on enzyme dimerization and not a specific amino acid residue [216]. In addition, Adolph *et al.* [216] found that A3Cs in dimer/monomer equilibrium, rather than stable dimers, constitute the favourable conformation for A3C to undergo

intersegment transfer. Altogether, these findings suggest mechanisms utilized by different A3s when they undergo intersegment transfer are different and the parameters required by an A3 to perform intersegment transfer could be better refined.

Due to the potential therapeutic value of disrupting the host-pathogen interaction thus fully unleashing the antiviral effects of A3s, the field has devoted a significant amount of work in defining the interacting residues between each individual A3 and Vif. It has been found that the interaction between Vif and each A3 enzyme (A3D/F, A3G and A3H) is different and that each pairwise combination can be mapped to physically distinct protein surfaces [284, 287, 292], suggesting each individual A3 likely exerts a unique type of selection pressure on Vif, in accordance with the observations that different A3s have very different biochemical properties. In good agreement with these previous studies, our work with the degradation-dependent inhibition mechanisms of A3G and A3H by Vif validated that the Vif interaction interface are different for single Z-domain A3H and double Z-domain A3G enzymes and that these different Vif-A3 interfaces affect the degradation efficiency of A3s by Vif [312]. Interestingly, our biochemical analysis suggests Vif-A3G interaction disrupts A3G dimerization and Vif interaction with A3H does not break its dimerization [312]. Because the double-domain A3G monomer is equivalent in size to the single-domain A3H dimer, it seemed likely that no large structural alterations within Cul5 E3-Vif-A3 would be required when an Cul5 E3-Vif mediates the degradation of double-domain as compared to single-domain A3 proteins and this could be an advantageous property for Vif as the degradation of A3s with different domain architecture would be more efficient. However, an optimal configuration still seems to be required by a Cul5 E3-Vif ligase complex to allow efficient degradation of A3s, based on the finding that despite A3G and A3H interacting with Vif_{HXB2} with similar affinities, the different Vif-A3 interfaces still enabled Vif_{HXB2} to degrade A3G 4-fold more than A3H [312]. It should be noted that all three recent studies [284, 287, 292] that allowed the identification of amino acid interfaces between Vif with A3G/F/H were performed using a similar approach involving identifying A3/Vif anchoring points using mutagenesis, functional infectivity assays and forced evolution experiments, followed by modeling of the interface using molecular docking [284, 292] or MD-simulation [287] based on pre-existing A3 structures solved by nuclear magnetic resonance (NMR) or crystallography. The limitations of this research are that they could not provide information on any potential structural rearrangements at the protein-protein interface upon an

A3 binding to Vif. Based on our research using purified proteins and rotational anisotropy, we demonstrated that a conformational change in A3G and A3H occurs upon Vif binding [312] (Fig. 5.2 and 5.3). Notwithstanding, solving the Cul5-E3-Vif-A3 co-crystal or NMR co-structure remains to be the "holy grail" in the A3 field. This would greatly strengthen our understanding of host-viral interaction and facilitate the small molecular inhibitor design that aims at inhibiting the A3-Vif interaction.

11.0 References

1. Kharsany, A.B. and Q.A. Karim, *HIV Infection and AIDS in Sub-Saharan Africa: Current Status, Challenges and Opportunities*. Open AIDS J, 2016. **10**: p. 34-48.
2. Barre-Sinoussi, F., A.L. Ross, and J.F. Delfraissy, *Past, present and future: 30 years of HIV research*. Nat Rev Microbiol, 2013. **11**(12): p. 877-83.
3. Peterlin, B.M. and D. Trono, *Hide, shield and strike back: how HIV-infected cells avoid immune eradication*. Nat Rev Immunol, 2003. **3**(2): p. 97-107.
4. Le Grice, S.F., *Human immunodeficiency virus reverse transcriptase: 25 years of research, drug discovery, and promise*. J Biol Chem, 2012. **287**(49): p. 40850-7.
5. Coffin, J.M., S.H. Hughes, and H.E. Varmus, *The Interactions of Retroviruses and their Hosts*, in *Retroviruses*, J.M. Coffin, S.H. Hughes, and H.E. Varmus, Editors. 1997: Cold Spring Harbor (NY).
6. Lu, K., X. Heng, and M.F. Summers, *Structural determinants and mechanism of HIV-1 genome packaging*. J Mol Biol, 2011. **410**(4): p. 609-33.
7. Vogt, V.M., *Retroviral Virions and Genomes*, in *Retroviruses*, J.M. Coffin, S.H. Hughes, and H.E. Varmus, Editors. 1997: Cold Spring Harbor (NY).
8. Malim, M.H. and M. Emerman, *HIV-1 accessory proteins--ensuring viral survival in a hostile environment*. Cell Host Microbe, 2008. **3**(6): p. 388-98.
9. Gramberg, T., N. Sunseri, and N.R. Landau, *Accessories to the crime: recent advances in HIV accessory protein biology*. Curr HIV/AIDS Rep, 2009. **6**(1): p. 36-42.
10. Collins, D.R. and K.L. Collins, *HIV-1 accessory proteins adapt cellular adaptors to facilitate immune evasion*. PLoS Pathog, 2014. **10**(1): p. e1003851.
11. Frankel, A.D. and J.A. Young, *HIV-1: fifteen proteins and an RNA*. Annu Rev Biochem, 1998. **67**: p. 1-25.
12. Sundquist, W.I. and H.G. Krausslich, *HIV-1 assembly, budding, and maturation*. Cold Spring Harb Perspect Med, 2012. **2**(7): p. a006924.
13. Arhel, N., *Revisiting HIV-1 uncoating*. Retrovirology, 2010. **7**: p. 96.
14. Dvorin, J.D. and M.H. Malim, *Intracellular trafficking of HIV-1 cores: journey to the center of the cell*. Curr Top Microbiol Immunol, 2003. **281**: p. 179-208.
15. Bukrinsky, M., *A hard way to the nucleus*. Mol Med, 2004. **10**(1-6): p. 1-5.
16. Goff, S.P., *Intracellular trafficking of retroviral genomes during the early phase of infection: viral exploitation of cellular pathways*. J Gene Med, 2001. **3**(6): p. 517-28.
17. Jacques, D.A., et al., *HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis*. Nature, 2016. **536**(7616): p. 349-53.
18. Hu, W.S. and S.H. Hughes, *HIV-1 reverse transcription*. Cold Spring Harb Perspect Med, 2012. **2**(10).
19. Sarafianos, S.G., et al., *Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition*. J Mol Biol, 2009. **385**(3): p. 693-713.
20. Arts, E.J. and D.J. Hazuda, *HIV-1 antiretroviral drug therapy*. Cold Spring Harb Perspect Med, 2012. **2**(4): p. a007161.
21. Harris, R.S., J.F. Hultquist, and D.T. Evans, *The restriction factors of human immunodeficiency virus*. J Biol Chem, 2012. **287**(49): p. 40875-83.
22. Zhang, H., et al., *The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA*. Nature, 2003. **424**(6944): p. 94-8.

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

24.   Harris, R.S., et al., *DNA deamination mediates innate immunity to retroviral infection*. Cell, 2003. **113**(6): p. 803-9.

25.   Sheehy, A.M., et al., *Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein*. Nature, 2002. **418**(6898): p. 646-50.

26.   Van Damme, N., et al., *The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein*. Cell Host Microbe, 2008. **3**(4): p. 245-52.

27.   Neil, S.J., T. Zang, and P.D. Bieniasz, *Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu*. Nature, 2008. **451**(7177): p. 425-30.

28.   Keckesova, Z., L.M. Ylinen, and G.J. Towers, *The human and African green monkey TRIM5alpha genes encode Ref1 and Lv1 retroviral restriction factor activities*. Proc Natl Acad Sci U S A, 2004. **101**(29): p. 10780-5.

29.   Stremlau, M., et al., *The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys*. Nature, 2004. **427**(6977): p. 848-53.

30.   Nisole, S., et al., *A Trim5-cyclophilin A fusion protein found in owl monkey kidney cells can restrict HIV-1*. Proc Natl Acad Sci U S A, 2004. **101**(36): p. 13324-8.

31.   Hrecka, K., et al., *Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein*. Nature, 2011. **474**(7353): p. 658-61.

32.   Laguette, N., et al., *SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx*. Nature, 2011. **474**(7353): p. 654-7.

33.   Goujon, C., et al., *Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection*. Nature, 2013. **502**(7472): p. 559-62.

34.   Hancks, D.C. and H.H. Kazazian, Jr., *Active human retrotransposons: variation and disease*. Curr Opin Genet Dev, 2012. **22**(3): p. 191-203.

35.   MacDuff, D.A., Z.L. Demorest, and R.S. Harris, *AID can restrict L1 retrotransposition suggesting a dual role in innate and adaptive immunity*. Nucleic Acids Res, 2009. **37**(6): p. 1854-67.

36.   Jern, P. and J.M. Coffin, *Host-retrovirus arms race: trimming the budget*. Cell Host Microbe, 2008. **4**(3): p. 196-7.

37.   LaRue, R.S., et al., *The artiodactyl APOBEC3 innate immune repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals*. BMC Mol Biol, 2008. **9**: p. 104.

38.   Munk, C., A. Willemsen, and I.G. Bravo, *An ancient history of gene duplications, fusions and losses in the evolution of APOBEC3 mutators in mammals*. BMC Evol Biol, 2012. **12**: p. 71.

39.   Jarmuz, A., et al., *An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22*. Genomics, 2002. **79**(3): p. 285-96.

40.   Esnault, C., et al., *APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses*. Nature, 2005. **433**(7024): p. 430-3.

41.   Bogerd, H.P., et al., *Cellular inhibitors of long interspersed element 1 and Alu retrotransposition*. Proc Natl Acad Sci U S A, 2006. **103**(23): p. 8780-5.

42.   Chiu, Y.L., et al., *High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition*. Proc Natl Acad Sci U S A, 2006. **103**(42): p. 15588-93.

43.   Jonsson, S.R., et al., *Evolutionarily conserved and non-conserved retrovirus restriction activities of artiodactyl APOBEC3F proteins*. Nucleic Acids Res, 2006. **34**(19): p. 5683-94.

44.   Armitage, A.E., et al., *Conserved footprints of APOBEC3G on Hypermutated human immunodeficiency virus type 1 and human endogenous retrovirus HERV-K(HML2) sequences*. J Virol, 2008. **82**(17): p. 8743-61.

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

46.   Khatua, A.K., et al., *Inhibition of LINE-1 and Alu retrotransposition by exosomes encapsidating APOBEC3G and APOBEC3F*. Virology, 2010. **400**(1): p. 68-75.

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

48.   Koyama, T., et al., *APOBEC3G oligomerization is associated with the inhibition of both Alu and LINE-1 retrotransposition*. PLoS One, 2013. **8**(12): p. e84228.

49.   Liddament, M.T., et al., *APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo*. Curr Biol, 2004. **14**(15): p. 1385-91.

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

51.   Zheng, Y.H., et al., *Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication*. J Virol, 2004. **78**(11): p. 6073-6.

52.   Dang, Y., et al., *Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family*. J Virol, 2006. **80**(21): p. 10522-33.

53.   Dang, Y., L.M. Siew, and Y.H. Zheng, *APOBEC3G is degraded by the proteasomal pathway in a Vif-dependent manner without being polyubiquitylated*. J Biol Chem, 2008. **283**(19): p. 13124-31.

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

55.   Lucifora, J., et al., *Specific and nonhepatotoxic degradation of nuclear hepatitis B virus cccDNA*. Science, 2014. **343**(6176): p. 1221-8.

56.   Hache, G., M.T. Liddament, and R.S. Harris, *The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain*. J Biol Chem, 2005. **280**(12): p. 10920-4.

57.   Navarro, F., et al., *Complementary function of the two catalytic domains of APOBEC3G*. Virology, 2005. **333**(2): p. 374-86.

58.   Bogerd, H.P., et al., *The intrinsic antiretroviral factor APOBEC3B contains two enzymatically active cytidine deaminase domains*. Virology, 2007. **364**(2): p. 486-93.

59.   Bonvin, M. and J. Greeve, *Effects of point mutations in the cytidine deaminase domains of APOBEC3B on replication and hypermutation of hepatitis B virus in vitro*. J Gen Virol, 2007. **88**(Pt 12): p. 3270-4.

60.   Shaw, G.M. and E. Hunter, *HIV transmission*. Cold Spring Harb Perspect Med, 2012. **2**(11).

61.   Xu, H., X. Wang, and R.S. Veazey, *Mucosal immunology of HIV infection*. Immunol Rev, 2013. **254**(1): p. 10-33.

62.   Rahm, N. and A. Telenti, *The role of tripartite motif family members in mediating susceptibility to HIV-1 infection*. Curr Opin HIV AIDS, 2012. **7**(2): p. 180-6.

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

64. Mariani, R., et al., *Species-specific exclusion of APOBEC3G from HIV-1 virions by Vif*. Cell, 2003. **114**(1): p. 21-31.

65. Compton, A.A. and M. Emerman, *Convergence and divergence in the evolution of the APOBEC3G-Vif interaction reveal ancient origins of simian immunodeficiency viruses*. PLoS Pathog, 2013. **9**(1): p. e1003135.

66. Feng, Y., et al., *Suppression of APOBEC3-mediated restriction of HIV-1 by Vif*. Front Microbiol, 2014. **5**: p. 450.

67. Etienne, L., et al., *Gene loss and adaptation to hominids underlie the ancient origin of HIV-1*. Cell Host Microbe, 2013. **14**(1): p. 85-92.

68. Jager, S., et al., *Vif hijacks CBF-beta to degrade APOBEC3G and promote HIV-1 infection*. Nature, 2012. **481**(7381): p. 371-5.

69. Zhang, W., et al., *T-cell differentiation factor CBF-beta regulates HIV-1 Vif-mediated evasion of host restriction*. Nature, 2012. **481**(7381): p. 376-9.

70. Yu, X., et al., *Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex*. Science, 2003. **302**(5647): p. 1056-60.

71. Kobayashi, M., et al., *Ubiquitination of APOBEC3G by an HIV-1 Vif-Cullin5-Elongin B-Elongin C complex is essential for Vif function*. J Biol Chem, 2005. **280**(19): p. 18573-8.

72. Xiao, Z., et al., *Assembly of HIV-1 Vif-Cul5 E3 ubiquitin ligase through a novel zinc-binding domain-stabilized hydrophobic interface in Vif*. Virology, 2006. **349**(2): p. 290-9.

73. Simon, V., et al., *Natural variation in Vif: differential impact on APOBEC3G/3F and a potential role in HIV-1 diversification*. PLoS Pathog, 2005. **1**(1): p. e6.

74. Kim, E.Y., et al., *Human APOBEC3G-mediated editing can promote HIV-1 sequence diversification and accelerate adaptation to selective pressure*. J Virol, 2010. **84**(19): p. 10402-5.

75. Armitage, A.E., et al., *APOBEC3G-induced hypermutation of human immunodeficiency virus type-1 is typically a discrete "all or nothing" phenomenon*. PLoS Genet, 2012. **8**(3): p. e1002550.

76. Delviks-Frankenberry, K.A., et al., *Minimal Contribution of APOBEC3-Induced G-to-A Hypermutation to HIV-1 Recombination and Genetic Variation*. PLoS Pathog, 2016. **12**(5): p. e1005646.

77. Berger, G., et al., *APOBEC3A is a specific inhibitor of the early phases of HIV-1 infection in myeloid cells*. PLoS Pathog, 2011. **7**(9): p. e1002221.

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

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

80. Refsland, E.W., J.F. Hultquist, and R.S. Harris, *Endogenous origins of HIV-1 G-to-A hypermutation and restriction in the nonpermissive T cell line CEM2n*. PLoS Pathog, 2012. **8**(7): p. e1002800.

81. Chaipan, C., et al., *APOBEC3G restricts HIV-1 to a greater extent than APOBEC3F and APOBEC3DE in human primary CD4+ T cells and macrophages*. J Virol, 2013. **87**(1): p. 444-53.

82.   Wittkopp, C.J., et al., *A Single Nucleotide Polymorphism in Human APOBEC3C Enhances Restriction of Lentiviruses*. PLoS Pathog, 2016. **12**(10): p. e1005865.

83.   Alce, T.M. and W. Popik, *APOBEC3G is incorporated into virus-like particles by a direct interaction with HIV-1 Gag nucleocapsid protein*. J Biol Chem, 2004. **279**(33): p. 34083-6.

84.   Cen, S., et al., *The interaction between HIV-1 Gag and APOBEC3G*. J Biol Chem, 2004. **279**(32): p. 33177-84.

85.   Douaisi, M., et al., *HIV-1 and MLV Gag proteins are sufficient to recruit APOBEC3G into virus-like particles*. Biochem Biophys Res Commun, 2004. **321**(3): p. 566-73.

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

87.   Burnett, A. and P. Spearman, *APOBEC3G multimers are recruited to the plasma membrane for packaging into human immunodeficiency virus type 1 virus-like particles in an RNA-dependent process requiring the NC basic linker*. J Virol, 2007. **81**(10): p. 5000-13.

88.   Khan, M.A., et al., *Analysis of the contribution of cellular and viral RNA to the packaging of APOBEC3G into HIV-1 virions*. Retrovirology, 2007. **4**: p. 48.

89.   Bach, D., et al., *Characterization of APOBEC3G binding to 7SL RNA*. Retrovirology, 2008. **5**: p. 54.

90.   Suspene, R., et al., *APOBEC3G is a single-stranded DNA cytidine deaminase and functions independently of HIV reverse transcriptase*. Nucleic Acids Res, 2004. **32**(8): p. 2421-9.

91.   Yu, Q., et al., *Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome*. Nat Struct Mol Biol, 2004. **11**(5): p. 435-42.

92.   Xu, H., et al., *Stoichiometry of the antiviral protein APOBEC3G in HIV-1 virions*. Virology, 2007. **360**(2): p. 247-56.

93.   Nowarski, R., et al., *Hypermutation by intersegmental transfer of APOBEC3G cytidine deaminase*. Nat Struct Mol Biol, 2008. **15**(10): p. 1059-66.

94.   Camaur, D. and D. Trono, *Characterization of human immunodeficiency virus type 1 Vif particle incorporation*. J Virol, 1996. **70**(9): p. 6106-11.

95.   Fouchier, R.A., et al., *Human immunodeficiency virus type 1 Vif does not influence expression or virion incorporation of gag-, pol-, and env-encoded proteins*. J Virol, 1996. **70**(12): p. 8263-9.

96.   Kozak, S.L., et al., *The anti-HIV-1 editing enzyme APOBEC3G binds HIV-1 RNA and messenger RNAs that shuttle between polysomes and stress granules*. J Biol Chem, 2006. **281**(39): p. 29105-19.

97.   Stopak, K.S., et al., *Distinct patterns of cytokine regulation of APOBEC3G expression and activity in primary lymphocytes, macrophages, and dendritic cells*. J Biol Chem, 2007. **282**(6): p. 3539-46.

98.   Gallois-Montbrun, S., et al., *Comparison of cellular ribonucleoprotein complexes associated with the APOBEC3F and APOBEC3G antiviral proteins*. J Virol, 2008. **82**(11): p. 5636-42.

99.   Refsland, E.W. and R.S. Harris, *The APOBEC3 family of retroelement restriction factors*. Curr Top Microbiol Immunol, 2013. **371**: p. 1-27.

100.   Ooms, M., et al., *HIV-1 Vif adaptation to human APOBEC3H haplotypes*. Cell Host Microbe, 2013. **14**(4): p. 411-21.

101.   Russell, R.A., et al., *APOBEC3G induces a hypermutation gradient: purifying selection at multiple steps during HIV-1 replication results in levels of G-to-A mutations that are high in DNA, intermediate in cellular viral RNA, and low in virion RNA*. Retrovirology, 2009. **6**: p. 16.

102.   Yang, B., et al., *Virion-associated uracil DNA glycosylase-2 and apurinic/apyrimidinic endonuclease are involved in the degradation of APOBEC3G-edited nascent HIV-1 DNA*. J Biol Chem, 2007. **282**(16): p. 11667-75.

103.   Kaiser, S.M. and M. Emerman, *Uracil DNA glycosylase is dispensable for human immunodeficiency virus type 1 replication and does not contribute to the antiviral effects of the cytidine deaminase Apobec3G*. J Virol, 2006. **80**(2): p. 875-82.

104.   Casartelli, N., et al., *The antiviral factor APOBEC3G improves CTL recognition of cultured HIV-infected T cells*. J Exp Med, 2010. **207**(1): p. 39-49.

105.   Norman, J.M., et al., *The antiviral factor APOBEC3G enhances the recognition of HIV-infected primary T cells by natural killer cells*. Nat Immunol, 2011. **12**(10): p. 975-83.

106.   Wood, N., et al., *HIV evolution in early infection: selection pressures, patterns of insertion and deletion, and the impact of APOBEC*. PLoS Pathog, 2009. **5**(5): p. e1000414.

107.   Monajemi, M., et al., *Emerging complexities of APOBEC3G action on immunity and viral fitness during HIV infection and treatment*. Retrovirology, 2012. **9**: p. 35.

108.   Squires, K.D., et al., *Impact of APOBEC Mutations on CD8+ T Cell Recognition of HIV Epitopes Varies Depending on the Restricting HLA*. J Acquir Immune Defic Syndr, 2015. **70**(2): p. 172-8.

109.   Monajemi, M., et al., *Positioning of APOBEC3G/F mutational hotspots in the human immunodeficiency virus genome favors reduced recognition by CD8+ T cells*. PLoS One, 2014. **9**(4): p. e93428.

110.   Jin, X., et al., *APOBEC3G/CEM15 (hA3G) mRNA levels associate inversely with human immunodeficiency virus viremia*. J Virol, 2005. **79**(17): p. 11513-6.

111.   Pace, C., et al., *Population level analysis of human immunodeficiency virus type 1 hypermutation and its relationship with APOBEC3G and vif genetic variation*. J Virol, 2006. **80**(18): p. 9259-69.

112.   Biasin, M., et al., *Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G: a possible role in the resistance to HIV of HIV-exposed seronegative individuals*. J Infect Dis, 2007. **195**(7): p. 960-4.

113.   Land, A.M., et al., *Human immunodeficiency virus (HIV) type 1 proviral hypermutation correlates with CD4 count in HIV-infected women from Kenya*. J Virol, 2008. **82**(16): p. 8172-82.

114.   Vazquez-Perez, J.A., et al., *APOBEC3G mRNA expression in exposed seronegative and early stage HIV infected individuals decreases with removal of exposure and with disease progression*. Retrovirology, 2009. **6**: p. 23.

115.   Albin, J.S. and R.S. Harris, *Interactions of host APOBEC3 restriction factors with HIV-1 in vivo: implications for therapeutics*. Expert Rev Mol Med, 2010. **12**: p. e4.

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

117. Desimmie, B.A., et al., *APOBEC3 proteins can copackage and comutate HIV-1 genomes*. Nucleic Acids Res, 2016. **44**(16): p. 7848-65.

118. Mulder, L.C., A. Harari, and V. Simon, *Cytidine deamination induced HIV-1 drug resistance*. Proc Natl Acad Sci U S A, 2008. **105**(14): p. 5501-6.

119. Bruner, K.M., et al., *Defective proviruses rapidly accumulate during acute HIV-1 infection*. Nat Med, 2016. **22**(9): p. 1043-9.

120. Coffin, J. and R. Swanstrom, *HIV pathogenesis: dynamics and genetics of viral populations and infected cells*. Cold Spring Harb Perspect Med, 2013. **3**(1): p. a012526.

121. Ji, J.P. and L.A. Loeb, *Fidelity of HIV-1 reverse transcriptase copying RNA in vitro*. Biochemistry, 1992. **31**(4): p. 954-8.

122. Rhodes, T.D., et al., *Genetic recombination of human immunodeficiency virus type 1 in one round of viral replication: effects of genetic distance, target cells, accessory genes, and lack of high negative interference in crossover events*. J Virol, 2005. **79**(3): p. 1666-77.

123. Rhodes, T., H. Wargo, and W.S. Hu, *High rates of human immunodeficiency virus type 1 recombination: near-random segregation of markers one kilobase apart in one round of viral replication*. J Virol, 2003. **77**(20): p. 11193-200.

124. Sadler, H.A., et al., *APOBEC3G contributes to HIV-1 variation through sublethal mutagenesis*. J Virol, 2010. **84**(14): p. 7396-404.

125. Fourati, S., et al., *Partially active HIV-1 Vif alleles facilitate viral escape from specific antiretrovirals*. AIDS, 2010. **24**(15): p. 2313-21.

126. Jern, P., et al., *Likely role of APOBEC3G-mediated G-to-A mutations in HIV-1 evolution and drug resistance*. PLoS Pathog, 2009. **5**(4): p. e1000367.

127. Noguera-Julian, M., et al., *Contribution of APOBEC3G/F activity to the development of low-abundance drug-resistant human immunodeficiency virus type 1 variants*. Clin Microbiol Infect, 2016. **22**(2): p. 191-200.

128. Ebrahimi, D., F. Anwar, and M.P. Davenport, *APOBEC3 has not left an evolutionary footprint on the HIV-1 genome*. J Virol, 2011. **85**(17): p. 9139-46.

129. Bogerd, H.P. and B.R. Cullen, *Single-stranded RNA facilitates nucleocapsid: APOBEC3G complex formation*. RNA, 2008. **14**(6): p. 1228-36.

130. Strebel, K. and M.A. Khan, *APOBEC3G encapsidation into HIV-1 virions: which RNA is it?* Retrovirology, 2008. **5**: p. 55.

131. Wang, T., et al., *Interaction with 7SL RNA but not with HIV-1 genomic RNA or P bodies is required for APOBEC3F virion packaging*. J Mol Biol, 2008. **375**(4): p. 1098-112.

132. Chelico, L., et al., *APOBEC3G DNA deaminase acts processively 3' --> 5' on single-stranded DNA*. Nat Struct Mol Biol, 2006. **13**(5): p. 392-9.

133. Ara, A., R.P. Love, and L. Chelico, *Different mutagenic potential of HIV-1 restriction factors APOBEC3G and APOBEC3F is determined by distinct single-stranded DNA scanning mechanisms*. PLoS Pathog, 2014. **10**(3): p. e1004024.

134. Wichroski, M.J., G.B. Robb, and T.M. Rana, *Human retroviral host restriction factors APOBEC3G and APOBEC3F localize to mRNA processing bodies*. PLoS Pathog, 2006. **2**(5): p. e41.

135. Soros, V.B., W. Yonemoto, and W.C. Greene, *Newly synthesized APOBEC3G is incorporated into HIV virions, inhibited by HIV RNA, and subsequently activated by RNase H*. PLoS Pathog, 2007. **3**(2): p. e15.

136. Wang, T., et al., *7SL RNA mediates virion packaging of the antiviral cytidine deaminase APOBEC3G*. J Virol, 2007. **81**(23): p. 13112-24.

137. Bulliard, Y., et al., *Functional analysis and structural modeling of human APOBEC3G reveal the role of evolutionarily conserved elements in the inhibition of human immunodeficiency virus type 1 infection and Alu transposition*. J Virol, 2009. **83**(23): p. 12611-21.

138. Huthoff, H., et al., *RNA-dependent oligomerization of APOBEC3G is required for restriction of HIV-1*. PLoS Pathog, 2009. **5**(3): p. e1000330.

139. Feng, Y. and L. Chelico, *Intensity of deoxycytidine deamination of HIV-1 proviral DNA by the retroviral restriction factor APOBEC3G is mediated by the noncatalytic domain*. J Biol Chem, 2011. **286**(13): p. 11415-26.

140. Chelico, L., et al., *Structural model for deoxycytidine deamination mechanisms of the HIV-1 inactivation enzyme APOBEC3G*. J Biol Chem, 2010. **285**(21): p. 16195-205.

141. Chelico, L., et al., *A model for oligomeric regulation of APOBEC3G cytosine deaminase-dependent restriction of HIV*. J Biol Chem, 2008. **283**(20): p. 13780-91.

142. Shlyakhtenko, L.S., et al., *Atomic force microscopy studies provide direct evidence for dimerization of the HIV restriction factor APOBEC3G*. J Biol Chem, 2011. **286**(5): p. 3387-95.

143. Goila-Gaur, R., et al., *Targeting APOBEC3A to the viral nucleoprotein complex confers antiviral activity*. Retrovirology, 2007. **4**: p. 61.

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

145. Berg, O.G., R.B. Winter, and P.H. von Hippel, *Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory*. Biochemistry, 1981. **20**(24): p. 6929-48.

146. von Hippel, P.H. and O.G. Berg, *Facilitated target location in biological systems*. J Biol Chem, 1989. **264**(2): p. 675-8.

147. Chelico, L., et al., *Biochemical basis of immunological and retroviral responses to DNA-targeted cytosine deamination by activation-induced cytidine deaminase and APOBEC3G*. J Biol Chem, 2009. **284**(41): p. 27761-5.

148. Halford, S.E. and J.F. Marko, *How do site-specific DNA-binding proteins find their targets?* Nucleic Acids Res, 2004. **32**(10): p. 3040-52.

149. Senavirathne, G., et al., *Single-stranded DNA scanning and deamination by APOBEC3G cytidine deaminase at single molecule resolution*. J Biol Chem, 2012. **287**(19): p. 15826-35.

150. Shlyakhtenko, L.S., et al., *Nanoscale structure and dynamics of ABOBEC3G complexes with single-stranded DNA*. Biochemistry, 2012. **51**(32): p. 6432-40.

151. Browne, E.P., C. Allers, and N.R. Landau, *Restriction of HIV-1 by APOBEC3G is cytidine deaminase-dependent*. Virology, 2009. **387**(2): p. 313-21.

152. Iwatani, Y., et al., *Deaminase-independent inhibition of HIV-1 reverse transcription by APOBEC3G*. Nucleic Acids Res, 2007. **35**(21): p. 7096-108.

153. Bishop, K.N., et al., *APOBEC3G inhibits elongation of HIV-1 reverse transcripts*. PLoS Pathog, 2008. **4**(12): p. e1000231.

154. Wang, X., et al., *The cellular antiviral protein APOBEC3G interacts with HIV-1 reverse transcriptase and inhibits its function during viral replication*. J Virol, 2012. **86**(7): p. 3777-86.

155. Adolph, M.B., J. Webb, and L. Chelico, *Retroviral restriction factor APOBEC3G delays the initiation of DNA synthesis by HIV-1 reverse transcriptase*. PLoS One, 2013. **8**(5): p. e64196.

156. Belanger, K., et al., *Binding of RNA by APOBEC3G controls deamination-independent restriction of retroviruses*. Nucleic Acids Res, 2013. **41**(15): p. 7438-52.

157. Gillick, K., et al., *Suppression of HIV-1 infection by APOBEC3 proteins in primary human CD4(+) T cells is associated with inhibition of processive reverse transcription as well as excessive cytidine deamination*. J Virol, 2013. **87**(3): p. 1508-17.

158. Guo, F., et al., *Inhibition of tRNA(3)(Lys)-primed reverse transcription by human APOBEC3G during human immunodeficiency virus type 1 replication*. J Virol, 2006. **80**(23): p. 11710-22.

159. Guo, F., et al., *The interaction of APOBEC3G with human immunodeficiency virus type 1 nucleocapsid inhibits tRNA3Lys annealing to viral RNA*. J Virol, 2007. **81**(20): p. 11322-31.

160. Luo, K., et al., *Cytidine deaminases APOBEC3G and APOBEC3F interact with human immunodeficiency virus type 1 integrase and inhibit proviral DNA formation*. J Virol, 2007. **81**(13): p. 7238-48.

161. Newman, E.N., et al., *Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity*. Curr Biol, 2005. **15**(2): p. 166-70.

162. Chaurasiya, K.R., et al., *Oligomerization transforms human APOBEC3G from an efficient enzyme to a slowly dissociating nucleic acid-binding protein*. Nat Chem, 2014. **6**(1): p. 28-33.

163. Mbisa, J.L., et al., *Human immunodeficiency virus type 1 cDNAs produced in the presence of APOBEC3G exhibit defects in plus-strand DNA transfer and integration*. J Virol, 2007. **81**(13): p. 7099-110.

164. Miyagi, E., et al., *Enzymatically active APOBEC3G is required for efficient inhibition of human immunodeficiency virus type 1*. J Virol, 2007. **81**(24): p. 13346-53.

165. Schumacher, A.J., et al., *The DNA deaminase activity of human APOBEC3G is required for Ty1, MusD, and human immunodeficiency virus type 1 restriction*. J Virol, 2008. **82**(6): p. 2652-60.

166. Liu, S., et al., *Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription*. Nat Struct Mol Biol, 2010. **17**(12): p. 1453-60.

167. Kobayashi, T., et al., *Quantification of deaminase activity-dependent and -independent restriction of HIV-1 replication mediated by APOBEC3F and APOBEC3G through experimental-mathematical investigation*. J Virol, 2014. **88**(10): p. 5881-7.

168. Pathak, V.K. and H.M. Temin, *Broad spectrum of in vivo forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: substitutions, frameshifts, and hypermutations*. Proc Natl Acad Sci U S A, 1990. **87**(16): p. 6019-23.

169. Li, Y., et al., *Molecular characterization of human immunodeficiency virus type 1 cloned directly from uncultured human brain tissue: identification of replication-competent and -defective viral genomes*. J Virol, 1991. **65**(8): p. 3973-85.

170. Vartanian, J.P., et al., *Selection, recombination, and G----A hypermutation of human immunodeficiency virus type 1 genomes*. J Virol, 1991. **65**(4): p. 1779-88.

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

172. Zennou, V. and P.D. Bieniasz, *Comparative analysis of the antiretroviral activity of APOBEC3G and APOBEC3F from primates*. Virology, 2006. **349**(1): p. 31-40.

173. Miyagi, E., et al., *Stably expressed APOBEC3F has negligible antiviral activity*. J Virol, 2010. **84**(21): p. 11067-75.

174. Mulder, L.C., et al., *Moderate influence of human APOBEC3F on HIV-1 replication in primary lymphocytes*. J Virol, 2010. **84**(18): p. 9613-7.

175. Russell, R.A., et al., *Distinct domains within APOBEC3G and APOBEC3F interact with separate regions of human immunodeficiency virus type 1 Vif*. J Virol, 2009. **83**(4): p. 1992-2003.

176. Song, C., et al., *Signals in APOBEC3F N-terminal and C-terminal deaminase domains each contribute to encapsidation in HIV-1 virions and are both required for HIV-1 restriction*. J Biol Chem, 2012. **287**(20): p. 16965-74.

177. Mbisa, J.L., W. Bu, and V.K. Pathak, *APOBEC3F and APOBEC3G inhibit HIV-1 DNA integration by different mechanisms*. J Virol, 2010. **84**(10): p. 5250-9.

178. Burdick, R.C., W.S. Hu, and V.K. Pathak, *Nuclear import of APOBEC3F-labeled HIV-1 preintegration complexes*. Proc Natl Acad Sci U S A, 2013. **110**(49): p. E4780-9.

179. Albin, J.S., et al., *Long-term restriction by APOBEC3F selects human immunodeficiency virus type 1 variants with restored Vif function*. J Virol, 2010. **84**(19): p. 10209-19.

180. Hache, G., et al., *Evolution of HIV-1 isolates that use a novel Vif-independent mechanism to resist restriction by human APOBEC3G*. Curr Biol, 2008. **18**(11): p. 819-24.

181. Hache, G., et al., *Optimal translation initiation enables Vif-deficient human immunodeficiency virus type 1 to escape restriction by APOBEC3G*. J Virol, 2009. **83**(11): p. 5956-60.

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

183. Feng, Y., R.P. Love, and L. Chelico, *HIV-1 viral infectivity factor (Vif) alters processive single-stranded DNA scanning of the retroviral restriction factor APOBEC3G*. J Biol Chem, 2013. **288**(9): p. 6083-94.

184. Langlois, M.A., et al., *Mutational comparison of the single-domained APOBEC3C and double-domained APOBEC3F/G anti-retroviral cytidine deaminases provides insight into their DNA target site specificities*. Nucleic Acids Res, 2005. **33**(6): p. 1913-23.

185. Kohli, R.M., et al., *A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase*. J Biol Chem, 2009. **284**(34): p. 22898-904.

186. Carpenter, M.A., et al., *Determinants of sequence-specificity within human AID and APOBEC3G*. DNA Repair (Amst), 2010. **9**(5): p. 579-87.

187. Rathore, A., et al., *The local dinucleotide preference of APOBEC3G can be altered from 5'-CC to 5'-TC by a single amino acid substitution*. J Mol Biol, 2013. **425**(22): p. 4442-54.

188. Bishop, K.N., R.K. Holmes, and M.H. Malim, *Antiviral potency of APOBEC proteins does not correlate with cytidine deamination*. J Virol, 2006. **80**(17): p. 8450-8.

189. Holmes, R.K., et al., *APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation*. *Comparisons with APOBEC3G*. J Biol Chem, 2007. **282**(4): p. 2587-95.

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

191.   Li, J., et al., *APOBEC3 multimerization correlates with HIV-1 packaging and restriction activity in living cells*. J Mol Biol, 2014. **426**(6): p. 1296-307.

192.   Abdel-Mohsen, M., et al., *Expression profile of host restriction factors in HIV-1 elite controllers*. Retrovirology, 2013. **10**: p. 106.

193.   Dang, Y., et al., *Identification of a single amino acid required for APOBEC3 antiretroviral cytidine deaminase activity*. J Virol, 2011. **85**(11): p. 5691-5.

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

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

196.   Genomes Project, C., et al., *A global reference for human genetic variation*. Nature, 2015. **526**(7571): p. 68-74.

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

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

199.   Sakurai, D., et al., *APOBEC3H polymorphisms associated with the susceptibility to HIV-1 infection and AIDS progression in Japanese*. Immunogenetics, 2015. **67**(4): p. 253-7.

200.   Naruse, T.K., et al., *APOBEC3H polymorphisms and susceptibility to HIV-1 infection in an Indian population*. J Hum Genet, 2016. **61**(3): p. 263-5.

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

202.   Starrett, G.J., et al., *The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis*. Nat Commun, 2016. **7**: p. 12918.

203.   Zhu, M., et al., *The eQTL-missense polymorphisms of APOBEC3H are associated with lung cancer risk in a Han Chinese population*. Sci Rep, 2015. **5**: p. 14969.

204.   Duggal, N.K., et al., *Identification and antiviral activity of common polymorphisms in the APOBEC3 locus in human populations*. Virology, 2013. **443**(2): p. 329-37.

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

206.   Love, R.P., H. Xu, and L. Chelico, *Biochemical analysis of hypermutation by the deoxycytidine deaminase APOBEC3A*. J Biol Chem, 2012. **287**(36): p. 30812-22.

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

208.   Logue, E.C., et al., *A DNA sequence recognition loop on APOBEC3A controls substrate specificity*. PLoS One, 2014. **9**(5): p. e97062.

209.   Shlyakhtenko, L.S., et al., *Interaction of APOBEC3A with DNA assessed by atomic force microscopy*. PLoS One, 2014. **9**(6): p. e99354.

210.   Feng, Y., et al., *Natural Polymorphisms and Oligomerization of Human APOBEC3H Contribute to Single-stranded DNA Scanning Ability*. J Biol Chem, 2015. **290**(45): p. 27188-203.

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

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

213. Mitra, M., et al., *Sequence and structural determinants of human APOBEC3H deaminase and anti-HIV-1 activities*. Retrovirology, 2015. **12**: p. 3.

214. Gu, J., et al., *Biochemical Characterization of APOBEC3H Variants: Implications for Their HIV-1 Restriction Activity and mC Modification*. J Mol Biol, 2016. **428**(23): p. 4626-4638.

215. Xiao, X., et al., *Crystal structures of APOBEC3G N-domain alone and its complex with DNA*. Nat Commun, 2016. **7**: p. 12193.

216. Adolph, M.B., et al., *Cytidine deaminase efficiency of the lentiviral viral restriction factor APOBEC3C correlates with dimerization*. Nucleic Acids Res, 2017.

217. Gabuzda, D.H., et al., *Role of vif in replication of human immunodeficiency virus type 1 in CD4+ T lymphocytes*. J Virol, 1992. **66**(11): p. 6489-95.

218. Blanc, D., et al., *Transcomplementation of VIF- HIV-1 mutants in CEM cells suggests that VIF affects late steps of the viral life cycle*. Virology, 1993. **193**(1): p. 186-92.

219. Sakai, H., et al., *Cell-dependent requirement of human immunodeficiency virus type 1 Vif protein for maturation of virus particles*. J Virol, 1993. **67**(3): p. 1663-6.

220. von Schwedler, U., et al., *Vif is crucial for human immunodeficiency virus type 1 proviral DNA synthesis in infected cells*. J Virol, 1993. **67**(8): p. 4945-55.

221. Madani, N. and D. Kabat, *An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral Vif protein*. J Virol, 1998. **72**(12): p. 10251-5.

222. Simon, J.H., et al., *Evidence for a newly discovered cellular anti-HIV-1 phenotype*. Nat Med, 1998. **4**(12): p. 1397-400.

223. Kao, S., et al., *The human immunodeficiency virus type 1 Vif protein reduces intracellular expression and inhibits packaging of APOBEC3G (CEM15), a cellular inhibitor of virus infectivity*. J Virol, 2003. **77**(21): p. 11398-407.

224. Martin, S.L., et al., *Trimeric structure for an essential protein in L1 retrotransposition*. Proc Natl Acad Sci U S A, 2003. **100**(24): p. 13815-20.

225. Stopak, K., et al., *HIV-1 Vif blocks the antiviral activity of APOBEC3G by impairing both its translation and intracellular stability*. Mol Cell, 2003. **12**(3): p. 591-601.

226. Piantadosi, A., et al., *Analysis of the percentage of human immunodeficiency virus type 1 sequences that are hypermutated and markers of disease progression in a longitudinal cohort, including one individual with a partially defective Vif*. J Virol, 2009. **83**(16): p. 7805-14.

227. Kieffer, T.L., et al., *G-->A hypermutation in protease and reverse transcriptase regions of human immunodeficiency virus type 1 residing in resting CD4+ T cells in vivo*. J Virol, 2005. **79**(3): p. 1975-80.

228. Vartanian, J.P., et al., *G-->A hypermutation of the human immunodeficiency virus type 1 genome: evidence for dCTP pool imbalance during reverse transcription*. Proc Natl Acad Sci U S A, 1994. **91**(8): p. 3092-6.

229. Krisko, J.F., et al., *APOBEC3G and APOBEC3F Act in Concert To Extinguish HIV-1 Replication*. J Virol, 2016. **90**(9): p. 4681-95.

230. Ara, A., et al., *Mechanism of enhanced HIV restriction by virion coencapsidated cytidine deaminases APOBEC3F and APOBEC3G*. J Virol, 2016.

231. Koning, F.A., et al., *Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets*. J Virol, 2009. **83**(18): p. 9474-85.

232. Refsland, E.W., et al., *Natural polymorphisms in human APOBEC3H and HIV-1 Vif combine in primary T lymphocytes to affect viral G-to-A mutation levels and infectivity.* PLoS Genet, 2014. **10**(11): p. e1004761.

233. Hultquist, J.F. and R.S. Harris, *Leveraging APOBEC3 proteins to alter the HIV mutation rate and combat AIDS.* Future Virol, 2009. **4**(6): p. 605.

234. Harris, R.S., *Enhancing immunity to HIV through APOBEC.* Nat Biotechnol, 2008. **26**(10): p. 1089-90.

235. Fisher, A.G., et al., *The sor gene of HIV-1 is required for efficient virus transmission in vitro.* Science, 1987. **237**(4817): p. 888-93.

236. Strebel, K., et al., *The HIV 'A' (sor) gene product is essential for virus infectivity.* Nature, 1987. **328**(6132): p. 728-30.

237. Hoglund, S., et al., *Role of vif during packing of the core of HIV-1.* Virology, 1994. **201**(2): p. 349-55.

238. Henriet, S., et al., *Vif is a RNA chaperone that could temporally regulate RNA dimerization and the early steps of HIV-1 reverse transcription.* Nucleic Acids Res, 2007. **35**(15): p. 5141-53.

239. Batisse, J., et al., *The role of Vif oligomerization and RNA chaperone activity in HIV-1 replication.* Virus Res, 2012. **169**(2): p. 361-76.

240. Binka, M., et al., *The activity spectrum of Vif from multiple HIV-1 subtypes against APOBEC3G, APOBEC3F, and APOBEC3H.* J Virol, 2012. **86**(1): p. 49-59.

241. Voon, D.C., Y.T. Hor, and Y. Ito, *The RUNX complex: reaching beyond haematopoiesis into immunity.* Immunology, 2015. **146**(4): p. 523-36.

242. Wang, H., et al., *Identification of HIV-1 Vif regions required for CBF-beta interaction and APOBEC3 suppression.* PLoS One, 2014. **9**(5): p. e95738.

243. Matsui, Y., et al., *Defining HIV-1 Vif residues that interact with CBFbeta by site-directed mutagenesis.* Virology, 2014. **449**: p. 82-7.

244. Zhou, X., et al., *Characterization of the interaction of full-length HIV-1 Vif protein with its key regulator CBFbeta and CRL5 E3 ubiquitin ligase components.* PLoS One, 2012. **7**(3): p. e33495.

245. Yang, X., J. Goncalves, and D. Gabuzda, *Phosphorylation of Vif and its role in HIV-1 replication.* J Biol Chem, 1996. **271**(17): p. 10121-9.

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

247. Mehle, A., et al., *Vif overcomes the innate antiviral activity of APOBEC3G by promoting its degradation in the ubiquitin-proteasome pathway.* J Biol Chem, 2004. **279**(9): p. 7792-8.

248. Luo, K., et al., *Primate lentiviral virion infectivity factors are substrate receptors that assemble with cullin 5-E3 ligase through a HCCH motif to suppress APOBEC3G.* Proc Natl Acad Sci U S A, 2005. **102**(32): p. 11444-9.

249. Xiao, Z., et al., *Zinc chelation inhibits HIV Vif activity and liberates antiviral function of the cytidine deaminase APOBEC3G.* FASEB J, 2007. **21**(1): p. 217-22.

250. Stanley, B.J., et al., *Structural insight into the human immunodeficiency virus Vif SOCS box and its role in human E3 ubiquitin ligase assembly.* J Virol, 2008. **82**(17): p. 8656-63.

251. Bergeron, J.R., et al., *The SOCS-box of HIV-1 Vif interacts with ElonginBC by induced-folding to recruit its Cul5-containing ubiquitin ligase complex.* PLoS Pathog, 2010. **6**(6): p. e1000925.

252. Wang, X., et al., *Interactions between HIV-1 Vif and human ElonginB-ElonginC are important for CBF-beta binding to Vif*. Retrovirology, 2013. **10**: p. 94.

253. Kamura, T., et al., *The Elongin BC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families*. Genes Dev, 1998. **12**(24): p. 3872-81.

254. Iwai, K., et al., *Identification of the von Hippel-lindau tumor-suppressor protein as part of an active E3 ubiquitin ligase complex*. Proc Natl Acad Sci U S A, 1999. **96**(22): p. 12436-41.

255. Kamura, T., et al., *VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases*. Genes Dev, 2004. **18**(24): p. 3055-65.

256. Guo, Y., et al., *Structural basis for hijacking CBF-beta and CUL5 E3 ligase complex by HIV-1 Vif*. Nature, 2014. **505**(7482): p. 229-33.

257. Fribourgh, J.L., et al., *Core binding factor beta plays a critical role by facilitating the assembly of the Vif-cullin 5 E3 ubiquitin ligase*. J Virol, 2014. **88**(6): p. 3309-19.

258. Kim, D.Y., et al., *CBFbeta stabilizes HIV Vif to counteract APOBEC3 at the expense of RUNX1 target gene expression*. Mol Cell, 2013. **49**(4): p. 632-44.

259. Hultquist, J.F., et al., *Vif proteins of human and simian immunodeficiency viruses require cellular CBFbeta to degrade APOBEC3 restriction factors*. J Virol, 2012. **86**(5): p. 2874-7.

260. Du, J., et al., *Differential requirements for HIV-1 Vif-mediated APOBEC3G degradation and RUNX1-mediated transcription by core binding factor beta*. J Virol, 2013. **87**(3): p. 1906-11.

261. Anderson, B.D. and R.S. Harris, *Transcriptional regulation of APOBEC3 antiviral immunity through the CBF-beta/RUNX axis*. Sci Adv, 2015. **1**(8): p. e1500296.

262. Yoshikawa, R., et al., *Small ruminant lentiviral Vif proteins commonly utilize cyclophilin A, an evolutionarily and structurally conserved protein, to degrade ovine and caprine APOBEC3 proteins*. Microbiol Immunol, 2016. **60**(6): p. 427-36.

263. Kane, J.R., et al., *Lineage-Specific Viral Hijacking of Non-canonical E3 Ubiquitin Ligase Cofactors in the Evolution of Vif Anti-APOBEC3 Activity*. Cell Rep, 2015. **11**(8): p. 1236-50.

264. Harris, R.S. and B.D. Anderson, *Evolutionary Paradigms from Ancient and Ongoing Conflicts between the Lentiviral Vif Protein and Mammalian APOBEC3 Enzymes*. PLoS Pathog, 2016. **12**(12): p. e1005958.

265. Aydin, H., M.W. Taylor, and J.E. Lee, *Structure-guided analysis of the human APOBEC3-HIV restrictome*. Structure, 2014. **22**(5): p. 668-84.

266. Albin, J.S., et al., *Dispersed sites of HIV Vif-dependent polyubiquitination in the DNA deaminase APOBEC3F*. J Mol Biol, 2013. **425**(7): p. 1172-82.

267. Huthoff, H. and M.H. Malim, *Identification of amino acid residues in APOBEC3G required for regulation by human immunodeficiency virus type 1 Vif and Virion encapsidation*. J Virol, 2007. **81**(8): p. 3807-15.

268. Mehle, A., et al., *Identification of an APOBEC3G binding site in human immunodeficiency virus type 1 Vif and inhibitors of Vif-APOBEC3G binding*. J Virol, 2007. **81**(23): p. 13235-41.

269. Russell, R.A. and V.K. Pathak, *Identification of two distinct human immunodeficiency virus type 1 Vif determinants critical for interactions with human APOBEC3G and APOBEC3F*. J Virol, 2007. **81**(15): p. 8201-10.

270. Yamashita, T., et al., *Identification of amino acid residues in HIV-1 Vif critical for binding and exclusion of APOBEC3G/F*. Microbes Infect, 2008. **10**(10-11): p. 1142-9.

271. Chen, G., et al., *A patch of positively charged amino acids surrounding the human immunodeficiency virus type 1 Vif SLVx4Yx9Y motif influences its interaction with APOBEC3G*. J Virol, 2009. **83**(17): p. 8674-82.

272. Dang, Y., et al., *Identification of a novel WxSLVK motif in the N terminus of human immunodeficiency virus and simian immunodeficiency virus Vif that is critical for APOBEC3G and APOBEC3F neutralization*. J Virol, 2009. **83**(17): p. 8544-52.

273. He, Z., et al., *Characterization of conserved motifs in HIV-1 Vif required for APOBEC3G and APOBEC3F interaction*. J Mol Biol, 2008. **381**(4): p. 1000-11.

274. Pery, E., et al., *Regulation of APOBEC3 proteins by a novel YXXL motif in human immunodeficiency virus type 1 Vif and simian immunodeficiency virus SIVagm Vif*. J Virol, 2009. **83**(5): p. 2374-81.

275. Kitamura, S., H. Ode, and Y. Iwatani, *Structural Features of Antiviral APOBEC3 Proteins are Linked to Their Functional Activities*. Front Microbiol, 2011. **2**: p. 258.

276. Ooms, M., et al., *The resistance of human APOBEC3H to HIV-1 NL4-3 molecular clone is determined by a single amino acid in Vif*. PLoS One, 2013. **8**(2): p. e57744.

277. Gourraud, P.A., et al., *APOBEC3H haplotypes and HIV-1 pro-viral vif DNA sequence diversity in early untreated human immunodeficiency virus-1 infection*. Hum Immunol, 2011. **72**(3): p. 207-12.

278. Bogerd, H.P., et al., *A single amino acid difference in the host APOBEC3G protein controls the primate species specificity of HIV type 1 virion infectivity factor*. Proc Natl Acad Sci U S A, 2004. **101**(11): p. 3770-4.

279. Mangeat, B., et al., *A single amino acid determinant governs the species-specific sensitivity of APOBEC3G to Vif action*. J Biol Chem, 2004. **279**(15): p. 14481-3.

280. Schrofelbauer, B., D. Chen, and N.R. Landau, *A single amino acid of APOBEC3G controls its species-specific interaction with virion infectivity factor (Vif)*. Proc Natl Acad Sci U S A, 2004. **101**(11): p. 3927-32.

281. Xu, H., et al., *A single amino acid substitution in human APOBEC3G antiretroviral enzyme confers resistance to HIV-1 virion infectivity factor-induced depletion*. Proc Natl Acad Sci U S A, 2004. **101**(15): p. 5652-7.

282. Letko, M., et al., *Vif proteins from diverse primate lentiviral lineages use the same binding site in APOBEC3G*. J Virol, 2013. **87**(21): p. 11861-71.

283. Lavens, D., et al., *Definition of the interacting interfaces of Apobec3G and HIV-1 Vif using MAPPIT mutagenesis analysis*. Nucleic Acids Res, 2010. **38**(6): p. 1902-12.

284. Letko, M., et al., *Identification of the HIV-1 Vif and Human APOBEC3G Protein Interface*. Cell Rep, 2015. **13**(9): p. 1789-99.

285. Smith, J.L. and V.K. Pathak, *Identification of specific determinants of human APOBEC3F, APOBEC3C, and APOBEC3DE and African green monkey APOBEC3F that interact with HIV-1 Vif*. J Virol, 2010. **84**(24): p. 12599-608.

286. Land, A.M., et al., *APOBEC3F determinants of HIV-1 Vif sensitivity*. J Virol, 2014. **88**(21): p. 12923-7.

287. Richards, C., et al., *The Binding Interface between Human APOBEC3F and HIV-1 Vif Elucidated by Genetic and Computational Approaches*. Cell Rep, 2015. **13**(9): p. 1781-8.

288. Albin, J.S., et al., *A single amino acid in human APOBEC3F alters susceptibility to HIV-1 Vif*. J Biol Chem, 2010. **285**(52): p. 40785-92.

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

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

291. Furukawa, A., et al., *Structure, interaction and real-time monitoring of the enzymatic reaction of wild-type APOBEC3G*. EMBO J, 2009. **28**(4): p. 440-51.

292. Ooms, M., M. Letko, and V. Simon, *The structural interface between HIV-1 Vif and human APOBEC3H*. J Virol, 2016.

293. Mercenne, G., et al., *HIV-1 Vif binds to APOBEC3G mRNA and inhibits its translation*. Nucleic Acids Res, 2010. **38**(2): p. 633-46.

294. Guerrero, S., et al., *Translational regulation of APOBEC3G mRNA by Vif requires its 5'UTR and contributes to restoring HIV-1 infectivity*. Sci Rep, 2016. **6**: p. 39507.

295. Marin, M., et al., *Human immunodeficiency virus type 1 Vif functionally interacts with diverse APOBEC3 cytidine deaminases and moves with them between cytoplasmic sites of mRNA metabolism*. J Virol, 2008. **82**(2): p. 987-98.

296. Goila-Gaur, R., et al., *HIV-1 Vif promotes the formation of high molecular mass APOBEC3G complexes*. Virology, 2008. **372**(1): p. 136-46.

297. Opi, S., et al., *Human immunodeficiency virus type 1 Vif inhibits packaging and antiviral activity of a degradation-resistant APOBEC3G variant*. J Virol, 2007. **81**(15): p. 8236-46.

298. Britan-Rosich, E., R. Nowarski, and M. Kotler, *Multifaceted counter-APOBEC3G mechanisms employed by HIV-1 Vif*. J Mol Biol, 2011. **410**(5): p. 1065-76.

299. Santa-Marta, M., et al., *HIV-1 Vif can directly inhibit apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G-mediated cytidine deamination by using a single amino acid interaction and without protein degradation*. J Biol Chem, 2005. **280**(10): p. 8765-75.

300. Greene, W.C., et al., *Novel targets for HIV therapy*. Antiviral Res, 2008. **80**(3): p. 251-65.

301. Nathans, R., et al., *Small-molecule inhibition of HIV-1 Vif*. Nat Biotechnol, 2008. **26**(10): p. 1187-92.

302. Cen, S., et al., *Small molecular compounds inhibit HIV-1 replication through specifically stabilizing APOBEC3G*. J Biol Chem, 2010. **285**(22): p. 16546-52.

303. Nowotny, B., et al., *Inducible APOBEC3G-Vif double stable cell line as a high-throughput screening platform to identify antiviral compounds*. Antimicrob Agents Chemother, 2010. **54**(1): p. 78-87.

304. Ejima, T., et al., *An anti-HIV-1 compound that increases steady-state expression of apoplipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G*. Int J Mol Med, 2011. **28**(4): p. 613-6.

305. Mohammed, I., et al., *SAR and Lead Optimization of an HIV-1 Vif-APOBEC3G Axis Inhibitor*. ACS Med Chem Lett, 2012. **3**(6): p. 465-469.

306. Matsui, M., et al., *Small molecules that inhibit Vif-induced degradation of APOBEC3G*. Virol J, 2014. **11**: p. 122.

307. Li, M., et al., *First-in-class small molecule inhibitors of the single-strand DNA cytosine deaminase APOBEC3G*. ACS Chem Biol, 2012. **7**(3): p. 506-17.

308. Olson, M.E., et al., *Small-molecule APOBEC3G DNA cytosine deaminase inhibitors based on a 4-amino-1,2,4-triazole-3-thiol scaffold*. ChemMedChem, 2013. **8**(1): p. 112-7.

309. Bennett, R.P., et al., *An analog of camptothecin inactive against Topoisomerase I is broadly neutralizing of HIV-1 through inhibition of Vif-dependent APOBEC3G degradation*. Antiviral Res, 2016. **136**: p. 51-59.

310. Greene, W.C., *A history of AIDS: looking back to see ahead*. Eur J Immunol, 2007. **37 Suppl 1**: p. S94-102.

311. Zuo, T., et al., *Small-molecule inhibition of human immunodeficiency virus type 1 replication by targeting the interaction between Vif and ElonginC*. J Virol, 2012. **86**(10): p. 5497-507.

312. Baig, T.T., Y. Feng, and L. Chelico, *Determinants of efficient degradation of APOBEC3 restriction factors by HIV-1 Vif*. J Virol, 2014. **88**(24): p. 14380-95.

313. Kazazian, H.H., Jr., *Mobile elements: drivers of genome evolution*. Science, 2004. **303**(5664): p. 1626-32.

314. Carmi, S., G.M. Church, and E.Y. Levanon, *Large-scale DNA editing of retrotransposons accelerates mammalian genome evolution*. Nat Commun, 2011. **2**: p. 519.

315. Beck, C.R., et al., *LINE-1 elements in structural variation and disease*. Annu Rev Genomics Hum Genet, 2011. **12**: p. 187-215.

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

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

318. Sanville, B., et al., *Adaptive evolution of Mus Apobec3 includes retroviral insertion and positive selection at two clusters of residues flanking the substrate groove*. PLoS Pathog, 2010. **6**: p. e1000974.

319. Harris, R.S. and J.P. Dudley, *APOBECs and virus restriction*. Virology, 2015. **479-480**: p. 131-45.

320. Koito, A. and T. Ikeda, *Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases*. Front Microbiol, 2013. **4**: p. 28.

321. Richardson, S.R., et al., *The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes*. Microbiol Spectr, 2015. **3**(2): p. MDNA3-0061-2014.

322. Cordaux, R. and M.A. Batzer, *The impact of retrotransposons on human genome evolution*. Nat Rev Genet, 2009. **10**(10): p. 691-703.

323. Beauregard, A., M.J. Curcio, and M. Belfort, *The take and give between retrotransposable elements and their hosts*. Annu Rev Genet, 2008. **42**: p. 587-617.

324. Lander, E.S., et al., *Initial sequencing and analysis of the human genome*. Nature, 2001. **409**(6822): p. 860-921.

325. Esnault, C., et al., *Dual inhibitory effects of APOBEC family proteins on retrotransposition of mammalian endogenous retroviruses*. Nucleic Acids Res, 2006. **34**(5): p. 1522-31.

326. Bogerd, H.P., et al., *APOBEC3A and APOBEC3B are potent inhibitors of LTR-retrotransposon function in human cells*. Nucleic Acids Res, 2006. **34**(1): p. 89-95.

327. Krokan, H.E. and M. Bjoras, *Base excision repair*. Cold Spring Harb Perspect Biol, 2013. **5**(4): p. a012583.

328. Anwar, F., M.P. Davenport, and D. Ebrahimi, *Footprint of APOBEC3 on the genome of human retroelements*. J Virol, 2013. **87**(14): p. 8195-204.

329. Jern, P., J.P. Stoye, and J.M. Coffin, *Role of APOBEC3 in genetic diversity among endogenous murine leukemia viruses*. PLoS Genet, 2007. **3**(10): p. 2014-22.

330. Jern, P. and J.M. Coffin, *Effects of retroviruses on host genome function*. Annu Rev Genet, 2008. **42**: p. 709-32.

331. Niewiadomska, A.M., et al., *Differential inhibition of long interspersed element 1 by APOBEC3 does not correlate with high-molecular-mass-complex formation or P-body association*. J Virol, 2007. **81**(17): p. 9577-83.

332. Kinomoto, M., et al., *All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition*. Nucleic Acids Res, 2007. **35**(9): p. 2955-64.

333. Stenglein, M.D. and R.S. Harris, *APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism*. J Biol Chem, 2006. **281**(25): p. 16837-41.

334. Muckenfuss, H., et al., *APOBEC3 proteins inhibit human LINE-1 retrotransposition*. J Biol Chem, 2006. **281**(31): p. 22161-72.

335. Horn, A.V., et al., *Human LINE-1 restriction by APOBEC3C is deaminase independent and mediated by an ORF1p interaction that affects LINE reverse transcriptase activity*. Nucleic Acids Res, 2014. **42**(1): p. 396-416.

336. Liang, W., et al., *APOBEC3DE Inhibits LINE-1 Retrotransposition by Interacting with ORF1p and Influencing LINE Reverse Transcriptase Activity*. PLoS One, 2016. **11**(7): p. e0157220.

337. Richardson, S.R., et al., *APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition*. Elife, 2014. **3**: p. e02008.

338. Rausch, J.W. and S.F. Le Grice, *'Binding, bending and bonding': polypurine tract-primed initiation of plus-strand DNA synthesis in human immunodeficiency virus*. Int J Biochem Cell Biol, 2004. **36**(9): p. 1752-66.

339. Hu, C., et al., *The HIV-1 central polypurine tract functions as a second line of defense against APOBEC3G/F*. J Virol, 2010. **84**(22): p. 11981-93.

340. Suspene, R., et al., *Twin gradients in APOBEC3 edited HIV-1 DNA reflect the dynamics of lentiviral replication*. Nucleic Acids Res, 2006. **34**(17): p. 4677-84.

341. Stanford, N.P., et al., *One- and three-dimensional pathways for proteins to reach specific DNA sites*. EMBO J, 2000. **19**(23): p. 6546-57.

342. Chelico, L., P. Pham, and M.F. Goodman, *Stochastic properties of processive cytidine DNA deaminases AID and APOBEC3G*. Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1517): p. 583-93.

343. Wedekind, J.E., et al., *Nanostructures of APOBEC3G support a hierarchical assembly model of high molecular mass ribonucleoprotein particles from dimeric subunits*. J Biol Chem, 2006. **281**(50): p. 38122-6.

344. Bennett, R.P., et al., *APOBEC3G subunits self-associate via the C-terminal deaminase domain*. J Biol Chem, 2008. **283**(48): p. 33329-36.

345. McDougall, W.M., C. Okany, and H.C. Smith, *Deaminase activity on single-stranded DNA (ssDNA) occurs in vitro when APOBEC3G cytidine deaminase forms homotetramers and higher-order complexes*. J Biol Chem, 2011. **286**(35): p. 30655-61.

346. Sheehy, A.M., N.C. Gaddis, and M.H. Malim, *The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif*. Nat Med, 2003. **9**(11): p. 1404-7.

347. Iwatani, Y., et al., *HIV-1 Vif-mediated ubiquitination/degradation of APOBEC3G involves four critical lysine residues in its C-terminal domain*. Proc Natl Acad Sci U S A, 2009. **106**(46): p. 19539-44.

348. Miller, J.H., V. Presnyak, and H.C. Smith, *The dimerization domain of HIV-1 viral infectivity factor Vif is required to block virion incorporation of APOBEC3G*. Retrovirology, 2007. **4**: p. 81.

349. Kao, S., et al., *Production of infectious human immunodeficiency virus type 1 does not require depletion of APOBEC3G from virus-producing cells*. Retrovirology, 2004. **1**: p. 27.

350. Karczewski, M.K. and K. Strebel, *Cytoskeleton association and virion incorporation of the human immunodeficiency virus type 1 Vif protein*. J Virol, 1996. **70**(1): p. 494-507.

351. Pillai, S.K., J.K. Wong, and J.D. Barbour, *Turning up the volume on mutational pressure: is more of a good thing always better? (A case study of HIV-1 Vif and APOBEC3)*. Retrovirology, 2008. **5**: p. 26.

352. Loeb, D.D., et al., *Complete mutagenesis of the HIV-1 protease*. Nature, 1989. **340**(6232): p. 397-400.

353. Cadima-Couto, I., et al., *HIV-1 Vif interaction with APOBEC3 deaminases and its characterization by a new sensitive assay*. J Neuroimmune Pharmacol, 2011. **6**(2): p. 296-307.

354. Bernacchi, S., et al., *RNA and DNA binding properties of HIV-1 Vif protein: a fluorescence study*. J Biol Chem, 2007. **282**(36): p. 26361-8.

355. Henriet, S., et al., *Tumultuous relationship between the human immunodeficiency virus type 1 viral infectivity factor (Vif) and the human APOBEC-3G and APOBEC-3F restriction factors*. Microbiol Mol Biol Rev, 2009. **73**(2): p. 211-32.

356. Uyttendaele, I., et al., *Random mutagenesis MAPPIT analysis identifies binding sites for Vif and Gag in both cytidine deaminase domains of Apobec3G*. PLoS One, 2012. **7**(9): p. e44143.

357. Nguyen, K.L., et al., *Codon optimization of the HIV-1 vpu and vif genes stabilizes their mRNA and allows for highly efficient Rev-independent expression*. Virology, 2004. **319**(2): p. 163-75.

358. Le Grice, S.F. and F. Gruninger-Leitch, *Rapid purification of homodimer and heterodimer HIV-1 reverse transcriptase by metal chelate affinity chromatography*. Eur J Biochem, 1990. **187**(2): p. 307-14.

359. Gao, F., et al., *The heterosexual human immunodeficiency virus type 1 epidemic in Thailand is caused by an intersubtype (A/E) recombinant of African origin*. J Virol, 1996. **70**(10): p. 7013-29.

360. Briggs, J.A., et al., *The stoichiometry of Gag protein in HIV-1*. Nat Struct Mol Biol, 2004. **11**(7): p. 672-5.

361. Zhu, P., et al., *Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions*. Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15812-7.

362. Creighton, S., L.B. Bloom, and M.F. Goodman, *Gel fidelity assay measuring nucleotide misinsertion, exonucleolytic proofreading, and lesion bypass efficiencies*. Methods Enzymol, 1995. **262**: p. 232-56.

363. Henriet, S., et al., *Cooperative and specific binding of Vif to the 5' region of HIV-1 genomic RNA*. J Mol Biol, 2005. **354**(1): p. 55-72.

364. Tian, C., et al., *Differential requirement for conserved tryptophans in human immunodeficiency virus type 1 Vif for the selective suppression of APOBEC3G and APOBEC3F*. J Virol, 2006. **80**(6): p. 3112-5.

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

366. Nakashima, M., et al., *Structural Insights into HIV-1 Vif-APOBEC3F Interaction*. J Virol, 2015. **90**(2): p. 1034-47.

367. Doehle, B.P., A. Schafer, and B.R. Cullen, *Human APOBEC3B is a potent inhibitor of HIV-1 infectivity and is resistant to HIV-1 Vif*. Virology, 2005. **339**(2): p. 281-8.

368. Salter, J.D., et al., *Core-binding factor beta increases the affinity between human Cullin 5 and HIV-1 Vif within an E3 ligase complex*. Biochemistry, 2012. **51**(44): p. 8702-4.

369. Iwabu, Y., et al., *Differential anti-APOBEC3G activity of HIV-1 Vif proteins derived from different subtypes*. J Biol Chem, 2010. **285**(46): p. 35350-8.

370. Jameson, D.M. and J.A. Ross, *Fluorescence polarization/anisotropy in diagnostics and imaging*. Chem Rev, 2010. **110**(5): p. 2685-708.

371. Wang, J., et al., *HIV-1 Vif promotes the G(1)- to S-phase cell-cycle transition*. Blood, 2011. **117**(4): p. 1260-9.

372. Halford, S.E., *An end to 40 years of mistakes in DNA-protein association kinetics?* Biochem Soc Trans, 2009. **37**(Pt 2): p. 343-8.

373. Prochnow, C., et al., *The APOBEC-2 crystal structure and functional implications for the deaminase AID*. Nature, 2007. **445**(7126): p. 447-51.

374. Lieberman, B.A. and S.K. Nordeen, *DNA intersegment transfer, how steroid receptors search for a target site*. J Biol Chem, 1997. **272**(2): p. 1061-8.

375. Luo, K., et al., *Amino-terminal region of the human immunodeficiency virus type 1 nucleocapsid is required for human APOBEC3G packaging*. J Virol, 2004. **78**(21): p. 11841-52.

376. Schafer, A., H.P. Bogerd, and B.R. Cullen, *Specific packaging of APOBEC3G into HIV-1 virions is mediated by the nucleocapsid domain of the gag polyprotein precursor*. Virology, 2004. **328**(2): p. 163-8.

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

378. Apolonia, L., et al., *Promiscuous RNA binding ensures effective encapsidation of APOBEC3 proteins by HIV-1*. PLoS Pathog, 2015. **11**(1): p. e1004609.

379. York, A., et al., *The RNA Binding Specificity of Human APOBEC3 Proteins Resembles That of HIV-1 Nucleocapsid*. PLoS Pathog, 2016. **12**(8): p. e1005833.

380. Mu, Y., et al., *A structural basis for the biochemical behavior of activation-induced deoxycytidine deaminase class-switch recombination-defective hyper-IgM-2 mutants*. J Biol Chem, 2012. **287**(33): p. 28007-16.

381. Shlyakhtenko, L.S., et al., *Atomic force microscopy studies of APOBEC3G oligomerization and dynamics*. J Struct Biol, 2013. **184**(2): p. 217-25.

382. Krzysiak, T.C., et al., *APOBEC2 is a monomer in solution: implications for APOBEC3G models*. Biochemistry, 2012. **51**(9): p. 2008-17.

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

384. Shandilya, S.M., et al., *Crystal structure of the APOBEC3G catalytic domain reveals potential oligomerization interfaces*. Structure, 2010. **18**(1): p. 28-38.

385. Bohn, M.F., et al., *The ssDNA Mutator APOBEC3A Is Regulated by Cooperative Dimerization*. Structure, 2015. **23**(5): p. 903-11.

386. King, J.J., et al., *Catalytic pocket inaccessibility of activation-induced cytidine deaminase is a safeguard against excessive mutagenic activity*. Structure, 2015. **23**(4): p. 615-27.

387. Shandilya, S.M., M.F. Bohn, and C.A. Schiffer, *A computational analysis of the structural determinants of APOBEC3's catalytic activity and vulnerability to HIV-1 Vif*. Virology, 2014. **471-473**: p. 105-16.

388. Lindic, N., et al., *Differential inhibition of LINE1 and LINE2 retrotransposition by vertebrate AID/APOBEC proteins*. Retrovirology, 2013. **10**: p. 156.

389. Wissing, S., et al., *Endogenous APOBEC3B restricts LINE-1 retrotransposition in transformed cells and human embryonic stem cells*. J Biol Chem, 2011. **286**(42): p. 36427-37.

390. Turelli, P., S. Vianin, and D. Trono, *The innate antiretroviral factor APOBEC3G does not affect human LINE-1 retrotransposition in a cell culture assay*. J Biol Chem, 2004. **279**(42): p. 43371-3.

391. Chen, H., et al., *APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons*. Curr Biol, 2006. **16**(5): p. 480-5.

392. Feschotte, C. and E.J. Pritham, *DNA transposons and the evolution of eukaryotic genomes*. Annu Rev Genet, 2007. **41**: p. 331-68.

393. Pace, J.K., 2nd and C. Feschotte, *The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage*. Genome Res, 2007. **17**(4): p. 422-32.

394. Brouha, B., et al., *Hot L1s account for the bulk of retrotransposition in the human population*. Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5280-5.

395. Dombroski, B.A., et al., *Isolation of an active human transposable element*. Science, 1991. **254**(5039): p. 1805-8.

396. Feng, Q., et al., *Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition*. Cell, 1996. **87**(5): p. 905-16.

397. Mathias, S.L., et al., *Reverse transcriptase encoded by a human transposable element*. Science, 1991. **254**(5039): p. 1808-10.

398. Piskareva, O. and V. Schmatchenko, *DNA polymerization by the reverse transcriptase of the human L1 retrotransposon on its own template in vitro*. FEBS Lett, 2006. **580**(2): p. 661-8.

399. Kurzynska-Kokorniak, A., et al., *DNA-directed DNA polymerase and strand displacement activity of the reverse transcriptase encoded by the R2 retrotransposon*. J Mol Biol, 2007. **374**(2): p. 322-33.

400. Zheng, L. and B. Shen, *Okazaki fragment maturation: nucleases take centre stage*. J Mol Cell Biol, 2011. **3**(1): p. 23-30.

401. Luan, D.D., et al., *Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition*. Cell, 1993. **72**(4): p. 595-605.

402. Cost, G.J., et al., *Human L1 element target-primed reverse transcription in vitro*. EMBO J, 2002. **21**(21): p. 5899-910.

403. Goodier, J.L. and H.H. Kazazian, Jr., *Retrotransposons revisited: the restraint and rehabilitation of parasites*. Cell, 2008. **135**(1): p. 23-35.

404. Yu, F., et al., *Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription*. Nucleic Acids Res, 2001. **29**(21): p. 4493-501.

405. Yoder, J.A., et al., *New 5' regions of the murine and human genes for DNA (cytosine-5)-methyltransferase*. J Biol Chem, 1996. **271**(49): p. 31092-7.

406. Yang, N. and H.H. Kazazian, Jr., *L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells*. Nat Struct Mol Biol, 2006. **13**(9): p. 763-71.

407. Malone, C.D. and G.J. Hannon, *Small RNAs as guardians of the genome*. Cell, 2009. **136**(4): p. 656-68.

408. Suzuki, J., et al., *Genetic evidence that the non-homologous end-joining repair pathway is involved in LINE retrotransposition*. PLoS Genet, 2009. **5**(4): p. e1000461.

409. Gasior, S.L., A.M. Roy-Engel, and P.L. Deininger, *ERCC1/XPF limits L1 retrotransposition*. DNA Repair (Amst), 2008. **7**(6): p. 983-9.

410. Stetson, D.B., et al., *Trex1 prevents cell-intrinsic initiation of autoimmunity*. Cell, 2008. **134**(4): p. 587-98.

411. Moldovan, J.B. and J.V. Moran, *The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition*. PLoS Genet, 2015. **11**(5): p. e1005121.

412. Goodier, J.L., L.E. Cheung, and H.H. Kazazian, Jr., *MOV10 RNA helicase is a potent inhibitor of retrotransposition in cells*. PLoS Genet, 2012. **8**(10): p. e1002941.

413. Zhang, A., et al., *RNase L restricts the mobility of engineered retrotransposons in cultured human cells*. Nucleic Acids Res, 2014. **42**(6): p. 3803-20.

414. Zhao, K., et al., *Modulation of LINE-1 and Alu/SVA retrotransposition by Aicardi-Goutieres syndrome-related SAMHD1*. Cell Rep, 2013. **4**(6): p. 1108-15.

415. Hu, S., et al., *SAMHD1 Inhibits LINE-1 Retrotransposition by Promoting Stress Granule Formation*. PLoS Genet, 2015. **11**(7): p. e1005367.

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

417. Conticello, S.G., et al., *DNA deamination in immunity: AID in the context of its APOBEC relatives*. Adv Immunol, 2007. **94**: p. 37-73.

418. Conticello, S.G., *The AID/APOBEC family of nucleic acid mutators*. Genome Biol, 2008. **9**(6): p. 229.

419. Rogozin, I.B., et al., *APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases predicted by computational analysis*. Cell Cycle, 2005. **4**(9): p. 1281-5.

420. Blanc, V. and N.O. Davidson, *APOBEC-1-mediated RNA editing*. Wiley Interdiscip Rev Syst Biol Med, 2010. **2**(5): p. 594-602.

421. Hulme, A.E., et al., *Selective inhibition of Alu retrotransposition by APOBEC3G*. Gene, 2007. **390**(1-2): p. 199-205.

422. Esnault, C., et al., *Restriction by APOBEC3 proteins of endogenous retroviruses with an extracellular life cycle: ex vivo effects and in vivo "traces" on the murine IAPE and human HERV-K elements*. Retrovirology, 2008. **5**: p. 75.

423. Lee, Y.N., M.H. Malim, and P.D. Bieniasz, *Hypermutation of an ancient human retrovirus by APOBEC3G*. J Virol, 2008. **82**(17): p. 8762-70.

424. Moran, J.V., et al., *High frequency retrotransposition in cultured mammalian cells*. Cell, 1996. **87**(5): p. 917-27.

425.   Kopera, H.C., et al., *LINE-1 Cultured Cell Retrotransposition Assay*. Methods Mol Biol, 2016. **1400**: p. 139-56.

426.   Hohjoh, H. and M.F. Singer, *Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA*. EMBO J, 1996. **15**(3): p. 630-9.

427.   Wei, W., et al., *Human L1 retrotransposition: cis preference versus trans complementation*. Mol Cell Biol, 2001. **21**(4): p. 1429-39.

428.   Wei, W., et al., *A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events*. Anal Biochem, 2000. **284**(2): p. 435-8.

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

430.   Carlow, D.C., S.A. Short, and R. Wolfenden, *Role of glutamate-104 in generating a transition state analogue inhibitor at the active site of cytidine deaminase*. Biochemistry, 1996. **35**(3): p. 948-54.

431.   Betts, L., et al., *Cytidine deaminase. The 2.3 A crystal structure of an enzyme: transition-state analog complex*. J Mol Biol, 1994. **235**(2): p. 635-56.

432.   Mol, C.D., et al., *Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: protein mimicry of DNA*. Cell, 1995. **82**(5): p. 701-8.

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

434.   Bulliard, Y., et al., *Structure-function analyses point to a polynucleotide-accommodating groove essential for APOBEC3A restriction activities*. J Virol, 2011. **85**(4): p. 1765-76.

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

436.   Nilsen, H., et al., *Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the UNG gene*. Nucleic Acids Res, 1997. **25**(4): p. 750-5.

437.   Karran, P., R. Cone, and E.C. Friedberg, *Specificity of the bacteriophage PBS2 induced inhibitor of uracil-DNA glycosylase*. Biochemistry, 1981. **20**(21): p. 6092-6.

438.   Langlois, M.A. and M.S. Neuberger, *Human APOBEC3G can restrict retroviral infection in avian cells and acts independently of both UNG and SMUG1*. J Virol, 2008. **82**(9): p. 4660-4.

439.   Lackey, L., et al., *Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination*. Cell Cycle, 2013. **12**(5): p. 762-72.

440.   Goodier, J.L., L.E. Cheung, and H.H. Kazazian, Jr., *Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition*. Nucleic Acids Res, 2013. **41**(15): p. 7401-19.

441.   Doucleff, M. and G.M. Clore, *Global jumping and domain-specific intersegment transfer between DNA cognate sites of the multidomain transcription factor Oct-1*. Proc Natl Acad Sci U S A, 2008. **105**(37): p. 13871-6.

442.   Hedglin, M., Y. Zhang, and P.J. O'Brien, *Isolating contributions from intersegmental transfer to DNA searching by alkyladenine DNA glycosylase*. J Biol Chem, 2013. **288**(34): p. 24550-9.

443.   Gorman, J., et al., *Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair*. Proc Natl Acad Sci U S A, 2012. **109**(45): p. E3074-83.

444.   Halford, S.E., *Hopping, jumping and looping by restriction enzymes*. Biochem Soc Trans, 2001. **29**(Pt 4): p. 363-74.

445. Terry, B.J., W.E. Jack, and P. Modrich, *Facilitated diffusion during catalysis by EcoRI endonuclease. Nonspecific interactions in EcoRI catalysis.* J Biol Chem, 1985. **260**(24): p. 13130-7.

446. Jeltsch, A., et al., *Pausing of the restriction endonuclease EcoRI during linear diffusion on DNA.* Biochemistry, 1994. **33**(34): p. 10215-9.